

1542  
(4-85)

U.S. DEPARTMENT OF COMMERCE  
PATENT AND TRADEMARK OFFICE

ONLINE SEARCH REQUEST FORM

8-393

\*\*\*\*\*  
USER Sheela Huff SERIAL NUMBER 08/23/85  
ART UNIT 1806 PHONE 30578666 DATE 8/26/96

Please give a detailed statement of requirements. Describe as specifically as possible the subject matter to be searched. Define any terms that may have special meaning. Give examples or relevant citations, authors, or keywords, if known.

You may include a copy of the broadest and or relevant claim(s).

Please search Seq. ID # 18, 2  
and 26, 27

TXS-

Sheela

Also Reverse transcribe AA seq's and search  
on NA databases.

\*\*\*\*\*  
STAFF USE ONLY

COMPLETED 8/22/96  
SEARCHER MAK  
ONLINE TIME 6 TOTAL TIME 10  
(in minutes)  
NO. OF DATABASES 6

SYSTEMS  
☐ CAS ONLINE  
☐ DARC/QUESTEL  
☐ DIALOG  
☐ SDC  
☒ OTHER MPRACT

\*\*\*\*\*

AVAILABILITY

(TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MParch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Aug 27 08:32:33 1996; MasPar time 625.55 Seconds  
896.015 Million cell updates/sec

Tabular output not generated.

Title: >US-08-231-565A-1

Description: (1-1559) from US08031565A.seq

Perfect Score: 1559

N.A. Sequence:

Comp: 1 AGCAGACAGAGACTCTCAT.....AAAAAAAAAAAAAAAAAAAA 1559

TCGCTCTCTCTCTGAGAGTA.....TTTTTTTTTTTTTTTTTTTT

Scoring table:

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 514334 seqs, 179763086 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

EST-STS  
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8  
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14  
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20  
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26  
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32  
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38  
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44  
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50  
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56  
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62  
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68  
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74  
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80  
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86  
87:EST87 88:EST88 89:EST89 90:EST90 91:STS1 92:STS2  
93:STS3 94:STS4 95:STS5 96:STS6 97:STS7  
EST-STS-TWO  
98:EST91 99:EST92 100:EST93 101:EST94 102:EST95  
103:EST96 104:EST97 105:EST98 106:EST99 107:EST100  
108:EST101 109:EST110 110:EST111 111:EST112 112:EST113  
113:EST114 114:EST115 115:EST116 116:EST117 117:EST118  
118:EST119 119:EST120 120:EST121 121:EST122 122:EST123  
123:EST124 124:EST125 125:EST126 126:EST127  
127:EST128 128:EST129 129:EST130 130:EST131 131:EST132

Database:

EST-STS-TWO  
98:EST91 99:EST92 100:EST93 101:EST94 102:EST95  
103:EST96 104:EST97 105:EST98 106:EST99 107:EST100  
108:EST101 109:EST110 110:EST111 111:EST112 112:EST113  
113:EST114 114:EST115 115:EST116 116:EST117 117:EST118  
118:EST119 119:EST120 120:EST121 121:EST122 122:EST123  
123:EST124 124:EST125 125:EST126 126:EST127  
127:EST128 128:EST129 129:EST130 130:EST131 131:EST132

132:enEST7 133:enEST8 134:enEST9 135:enEST10 136:enEST11  
137:enEST12

Statistics: Mean 11.935; Variance 4.909; scale 2.431

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description	Pred. No.
c	1	519	33.3	642	134	HS587267	yx91e07.s1	Homo sapie	0.00e+00
c	2	519	33.3	642	114	N26587	yx91e07.s1	Homo sapie	0.00e+00
c	3	488	31.3	593	113	N23770	yx35a12.r1	Homo sapie	0.00e+00
c	4	488	31.3	593	136	HS770262	yx35a12.r1	Homo sapie	0.00e+00
c	5	482	30.9	522	134	HS510255	yx11h07.s1	Homo sapie	0.00e+00
c	6	482	30.9	522	111	H97510	yx11h07.s1	Homo sapie	0.00e+00
c	7	468	30.0	540	111	HS7498	yx11f07.s1	Homo sapie	0.00e+00
c	8	468	30.0	540	133	HS498253	yx11f07.s1	Homo sapie	0.00e+00
c	9	441	28.3	473	112	N20091	yx36a12.s1	Homo sapie	0.00e+00
c	10	441	28.3	473	129	HS091255	yx36a12.s1	Homo sapie	0.00e+00
c	11	423	27.1	433	113	N24677	yx91a09.s1	Homo sapie	0.00e+00
c	12	423	27.1	433	135	HS677260	yx91a09.s1	Homo sapie	0.00e+00
c	13	394	25.3	396	136	HS796259	yx97f10.s1	Homo sapie	0.00e+00
c	14	394	25.3	396	113	N24796	yx97f10.s1	Homo sapie	0.00e+00
c	15	392	25.1	401	114	N26562	yx62h05.s1	Homo sapie	0.00e+00
c	16	392	25.1	401	134	HS366267	yx62h05.s1	Homo sapie	0.00e+00
c	17	357	22.9	365	118	N39656	yx91e07.r1	Homo sapie	0.00e+00
c	18	357	22.9	365	135	HS656272	yx91e07.r1	Homo sapie	0.00e+00
c	19	357	22.9	370	117	N35680	yx62h05.r1	Homo sapie	0.00e+00
c	20	357	22.9	370	135	HS680270	yx62h05.r1	Homo sapie	0.00e+00
c	21	353	22.6	393	135	HS656271	yx91a09.r1	Homo sapie	0.00e+00
c	22	353	22.6	393	117	N36656	yx91a09.r1	Homo sapie	0.00e+00
c	23	289	18.5	303	82	T27884	EST19370	Homo sapiens	1.28e-249
c	24	224	14.4	406	131	HS298280	yx23g10.s1	Homo sapie	4.48e-184
c	25	223	14.3	226	130	HS198272	yx25h10.s1	Homo sapie	4.48e-183
c	26	223	14.3	226	116	N32198	yx25h10.s1	Homo sapie	4.48e-177
c	27	217	13.9	349	79	T10292	seq1461	Homo sapiens	4.52e-177
c	28	216	13.9	505	23	H64579	yx63h07.s1	Homo sapie	4.52e-176
c	29	215	13.8	370	44	M77888	EST01472	Homo sapiens	4.52e-175
c	30	215	13.8	387	44	M77964	EST01548	Homo sapiens	4.52e-175
c	31	215	13.8	445	130	HS159264	yx94f02.s1	Homo sapie	4.52e-175
c	32	215	13.8	445	114	N26159	yx94f02.s1	Homo sapie	4.52e-175
c	33	215	13.7	461	11	H24331	ym51e07.s1	Homo sapie	4.52e-175
c	34	213	13.7	439	63	R61887	yh10g05.s1	Homo sapie	4.50e-173
c	35	214	13.7	443	5	H05066	yl69f06.s1	Homo sapie	4.51e-174
c	36	213	13.7	453	13	H29593	ym61b01.s1	Homo sapie	4.50e-173
c	37	214	13.7	463	130	HS135267	yy20d04.s1	Homo sapie	4.51e-174
c	38	214	13.7	463	116	N35135	yy20d04.s1	Homo sapie	4.51e-174
c	39	212	13.6	349	81	T24036	seq2194	Homo sapiens	4.49e-172
c	40	212	13.6	349	77	T03928	seq2491	Homo sapiens	4.49e-172
c	41	212	13.6	441	72	R92703	yc03c12.s1	Homo sapie	4.49e-172
c	42	212	13.6	495	61	R53536	yg84c08.s1	Homo sapie	4.49e-172
c	43	210	13.5	452	20	H53109	yg84c08.s1	Homo sapie	4.46e-170
c	44	210	13.5	499	82	T26553	AB31486F	Homo sapiens	4.46e-170
c	45	210	13.5	558	134	HS562300	yy63e01.s1	Homo sapie	4.46e-170

## ALIGNMENTS

RESULT 1

ID H587267 standard; RNA; EST; 642 BP.  
 AC N26587;  
 DT 30-DEC-1995 (Rel. 46, Created)  
 DT 30-DEC-1995 (Rel. 46, Last updated, Version 1)  
 DE yx9le07.s1 Homo sapiens cDNA clone 269124 3' similar to contains  
 DE Alu repetitive element;  
 KW EST.  
 OS Homo sapiens (human)  
 OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;  
 OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.  
 RN [1]  
 RP 1-642  
 RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,  
 RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,  
 RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,  
 RA Trevasakis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;  
 RT "The WashU-Merck EST Project";  
 RL Unpublished.  
 CC Contact: Wilson RK WashU-Merck EST Project Washington University  
 CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,  
 CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:  
 CC est@watson.wustl.edu High quality sequence stops: 385 Source: IMAGE  
 CC Consortium, LLNL This clone is available royalty-free through LLNL  
 CC ; contact the IMAGE Consortium (info@image.llnl.gov) for further  
 CC information. NCBI gi: 1140935  
 FH Key Location/Qualifiers  
 FT source  
 FT 1..642  
 FT /organism="Homo sapiens"  
 FT /clone="269124"  
 FT /note="human"  
 FT mRNA  
 FT <1..>642  
 SQ Sequence 642 BP; 182 A; 135 C; 152 G; 166 T; 7 other;  
 Query Match 33.3%; Score 519; DB 134; Length 642;  
 Best Local Similarity 97.8%; Pred. No. 0.00e+00;  
 Matches 544; Conservative 0; Mismatches 9; Indels 3; Gaps 3;  
 Db 1 gaatctctgacattatgattctctgacattctgtgagaggatttgggaacc-a 59  
 Cp 1522 GATCTCTGATCTTTATGATTTCTTCGACATTTCTGACAGGGATTGGGACCCA 1463  
 Db 60 agagagctaaattgcataggtacatttaaggcagtcacagggcagcactaaggcag 119  
 Cp 1462 AGAGAGCTAAATTGCCATAGGTACATTTAAGGCGAGTACACAGGCTACGACTAAGGCAG 1403  
 Db 120 tacatagtagctttactatttaaatcccttaccattcttggttagtagaataac 179  
 Cp 1402 TACATAGTAGCTTTACTATTATTAATCCTTCTACCAATTTCTGGTTAGTAGAATAATAC 1343  
 Db 180 ttgtcattagtagatgacattgagcttgaagtgaatgaatgactgogttatat 239  
 Cp 1342 TTGCTATTAGTAGATGAGTTCAGCCCTTCAGTCAATCTAATAGACTGCGCTATAT 1283  
 Db 240 gtcttacctaatagatagatccagccagggtggtggctcaqgc-tgtattccagcac 298  
 Cp 1282 GTCTTACTAGATATAGATATCAGCGAGGGGTGGTGGCTCGACCTGTAAATTCAGACAC 1223  
 Db 299 ttggaggctgagggcgagcatcacctcaggtcaggagtttgagaccagcctgaccaa 358  
 Cp 1222 TTGGGAGGCTGAGGCGGAGCATCACCTTGAGTCAGAGTTTGAACACGCGCTGACCAA 1163  
 Db 359 catggagaaccctctctactaaactacaaattagtcaggcatagtgccacnct 418  
 |||||||

Cp 1162 CATGAGAAACCCCGTCTCTACTTAAACTACAAAATAGTACGCATAGTGGCGCAGCC 1103  
 Db 419 gttLaatcccaagctactcagaggaggaagcaggaattgctgaacctgggagcgag 478  
 |||||||  
 Cp 1102 TGTAAATCCCACTACTACTAGGAGGCTAAGCGAGGAATGCTTGACCTGGGAGCGGAG 1043  
 |||||||  
 Db 479 gttatggtgagcaagatcgccattgcactccagctggcgcaaaaagcgaactcc 538  
 |||||||  
 Cp 1042 GTTATGG-TGAGCCAGATCGGCCATTGCACTCCAGCTGGGCAACAAAGCGCAACTC 984  
 |||||||  
 Db 539 catctcaaaaaaagag 554  
 |||||||  
 Cp 983 CATCTCAAAAAAAG 968  
 |||||||  
 RESULT 2  
 LOCUS N26587 642 bp mRNA EST 29-DEC-1995  
 DEFINITION yx9le07.s1 Homo sapiens cDNA clone 269124 3' similar to contains  
 Alu repetitive element;  
 ACCESSION N26587  
 NID g1140935  
 KEYWORDS EST.  
 SOURCE human clone=269124 primer=ml3-40 forward library=Soares melanocyte  
 2NBHM vector=pf7T30 (Pharmacia) with a modified polylinker  
 host=DH10B (ampicillin resistant) Reitel=Not I Rsite2=Eco RI Male.  
 1st strand cDNA was primed with a Not I - oligo(dT) primer  
 [5'-TGTTACCAATCTGAAGTGGCGGCGGCGGAGTTTCTTTTCTTTT-3'],  
 double-stranded cDNA was size selected, ligated to Eco RI adapters  
 (Pharmacia), digested with Not I and cloned into the Not I and Eco  
 RI sites of a modified pf7T3 vector (Pharmacia). Library  
 constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal  
 foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.  
 Albino.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;  
 Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
 Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;  
 Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.  
 REFERENCE 1 (bases 1 to 642)  
 AUTHORS Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,  
 Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,  
 Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,  
 Trevasakis E., Waterston R., Williamson A., Wohlmann P., and  
 Wilson R.  
 TITLE The WashU-Merck EST Project  
 JOURNAL Unpublished (1995)  
 COMMENT  
 Contact: Wilson RK  
 WashU-Merck EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 High quality sequence stops: 385  
 Source: IMAGE Consortium, LLNL  
 This clone is available royalty-free through LLNL ; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 NCBI gi: 1140935 Location/Qualifiers  
 1..642  
 /organism="Homo sapiens"  
 /clone="269124"  
 FEATURES  
 source

Aug 27 08:30

US-08-231-565A-1.rst

5

```

/Note="human"
BASE COUNT 182 a 135 c 152 g 166 t 7 others
ORIGIN
Query Match 33.3%; Score 519; DB 114; Length 642;
Best Local Similarity 97.8%; Pred. No. 0.00e+00;
Matches 544; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

Db 1 gaatctgacatttatttcttctgcacattcttgcagagggatttgggaacc-a 59
Cp 1522 GAATCTCTGATCCTTTATGATTTCTTCGCACATTTCTGTCAGAGGAGATTGGGAACCA 1463
Db 60 agagagctaaattgcacatgtagtattacatttaagcagtcacacagggatcagcactaagcgag 119
Cp 1462 AGAGAGCTAAATTGCGATGAGTACATTAAGCAGCAGTACACAGGATCAGCAGCTAAGCGAG 1403
Db 120 tacatagtagcttttacttatttatttaccattcttaccatttctggttttagtagaataac 179
Cp 1402 TACATAGTAGCTTTTACTTATTAAATCCTTACCAATTTCTGCTTTAGTAGAATAATAC 1343
Db 180 ttgtcattagtagaataagcattgagcttgaagtgaatgaaatgaaatgactgacttatat 239
Cp 1342 TTGTCATTAGTAGAATAGCATTGAGCCTTGAAGTGAATGAAATGAAATGAAATGAAATGAAAT 1283
Db 240 gtcttaactaagatatagtagatccagcagggcgtggtggtcaagc-tgtaattccagcac 298
Cp 1282 GTCTTACCTTAAGATATAGATCAGCAGCGGCGTGGCTCAGCCTGTAATTCAGCAC 1223
Db 299 ttggaggctgagggcgagcatcacctgaggtcaggagtttgagaccagcctgaccaa 358
Cp 1222 TTGGGAGGCTGAGCGGGCAGATACCTGAGGTGAGGAGTTGACACAGCCTGACCAA 1163
Db 359 catggaagaacccctctctactaaactacaaattagtcaggcagcatagtcgacacnct 418
Cp 1162 CATGGAAGAACCCGCTCTTACTTAAATACAAATAGTACAGGATAGTGGCGCAGCC 1103
Db 419 gttatccacagtaactcaggagctaaagcagaggaattgcttgaacctgggagcgag 478
Cp 1102 TGTAATCCAGCTTACTCAGGAGCTTAAGCAGAGGAATTTGCTTCACTGGGAGCGGAG 1043
Db 479 gttatgggtgagccaagatcgccattgcaactccagcctgggcaaaaagcgaactcc 538
Cp 1042 GTTATGG-TGAGCGAAGATCGCCCATTTGCATTCAGGCTGGGCAACAAAGCAAACTC 984
Db 539 catctcaaaaaaagag 554
Cp 983 CATCTCAAAAAAAG 968

RESULT 3
LOCUS N23770 593 bp mRNA EST 28-DEC-1995
DEFINITION YK35al2.r1 Homo sapiens cDNA clone 263710 5'.
ACCESSION N23770
NID g1137920
KEYWORDS EST.
SOURCE human clone=263710 primer=T7 library=Soares melanocyte 2Nbm
vector=pT73D (Pharmacia) with a modified polylinker host=DH10B
(ampicillin resistant) Reite1=Not I Reite2=Eco RI Male. 1st strand
cDNA was primed with a Not I - oligo (dT) primer
[5'-TCTTACCAATCTGAAGTGGGAGCCGCCGCTTTTCTTTTCTTTTCTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT73 vector (Pharmacia). Library

```

Aug 27 08:30

US-08-231-565A-1.rst

6

```

constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.
Albino.
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 593)
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfs, T., Soares, M., Tan, F.,
Trevaak, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 414
Source: IMAGE Consortium, LIND
This clone is available royalty-free through LIND; contact the
IMAGE Consortium (info@image.lind.gov) for further information.
NCBI gi: 1137920 Location/Qualifiers
source 1..593
/organism="Homo sapiens"
/clone="263710"
/note="human"
BASE COUNT 165 a 140 c 137 g 146 t 5 others
ORIGIN
Query Match 31.3%; Score 488; DB 113; Length 593;
Best Local Similarity 96.5%; Pred. No. 0.00e+00;
Matches 555; Conservative 0; Mismatches 10; Indels 10; Gaps 10;
Db 2 agcagacagagagactctcatttaaggaaggtgctgctgacccctacaagatgccaa 61
Qy 1 AGCAGACAGAGAGACTCTCATTAAGGAAGGTGCTGCTGCTGACCTACAGATGCCAA 60
Db 62 gagaagatgctcactctatctatggttaccaccaagaggcgccactcttacacca 121
Qy 61 GAGAGATGCTCACTTCTATCTATGTTACCCCAAGAGGGCGGCGGCTTTACACCA 120
Db 122 cgctgaagagggccgtggatcgccatcctgacagtgcctgggagcttactgctca 181
Qy 121 CGCTGAAGAGGGCGGCTGGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCG 180
Db 182 tcgctgttggtatgtagaagacgaataatgatacagagccttgatgataaagcttc 241
Qy 181 TCGGCTGTGTTATGTTAGAACGCAATATGATACAGAGCCTTGCATGATAAAGCTTC 240
Db 242 atgtggcactcaatgctcttaacaagaagatgcccaagaaggggttgcacatcg 301
Qy 241 ATGTGGCACTCAATGTGCTTTAACAGAAGATGCCCAAGAAGGGTTTGCATCACTGG 300
Db 302 acagcaagagtctcttccaagagaaaaactgtgaacctgtggttcccaatgctccactg 361

```



[illegible]

RESULT	4	
ID	HS770262	standard; RNA; EST; 593 BP.
AC	N23770;	
DT	29-DEC-1995 (Rel. 46, Created)	
DT	29-DEC-1995 (Rel. 46, Last updated, Version 1)	
DE	yx35a12.r1 Homo sapiens cDNA clone 263710 5'.	
KW	EST.	
OS	Homo sapiens (human)	
OC	Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;	
OC	Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.	
RN	[1]	
RP	1-593	
RA	Hillier L., Clark N., Dubouque T., Elliston K., Hawkins M.,	
RA	Holtman M., Holtman M., Kucaba T., Le M., Lennon G., Marra M.,	
RA	Parsons J., Rifkin L., Robling T., Soares M., Tan F.,	
RA	Trevaskis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;	
RT	"The WashU-Merck EST Project";	
RL	Unpublished.	
CC	Contact: Wilson RK WashU-Merck EST Project Washington University	
CC	School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,	
CC	MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:	
CC	est@watson.wustl.edu High quality sequence stops: 414 Source: IMAGE	
CC	Consortium, LINL This clone is available royalty-free through LINL	
CC	i contact the IMAGE Consortium ( <a href="mailto:info@image.llnl.gov">info@image.llnl.gov</a> ) for further	
CC	information. NCBI gi: 1137920	
FF	Key	Location/Qualifiers
FH	source	1..593
FT		/organism="Homo sapiens"
FT		/clone="263710"
FT		/note="human"
FT	mRNA	<1..>593
SQL	Sequence 593 BP; 165 A; 140 C; 137 G; 146 T; 5 other;	

	Query Match	31.3%;	Score 488;	DB 136;	Length 593;
	Best Local Similarity	96.5%;	Pred. No. 0.00e+00;		
	Matches 555;	Conservative	0;	Mismatches 10;	Indels 10;
	Gaps	1			
Db	2	agcagacagaggactctcattaaaggaggtgtccttgccctgacctacaagatgccaa	61		
Qy	1	ACGACACAGAGGACTCTCATTAAAGGAAGGTGTCTGTGGCCTGACCCCTACAGATGCCAA	60		
Db	62	gagaagatgctcacttcctatctatggttaccaccaagaaggggcaggccactcttacacca	121		

61	GAGAGAGTCTCAC	TTCA	TCTAT	TGGTTACCC	CAAGAGGGCCACGGCCACTCTTTACACCA	121	
122	cggtggaagagccgc	gggatcg	gcctcatc	gcagctgatc	ctctgggagctcttactgctca	181	
121	CGGCTGAAGAGGCC	CTGGGATCG	CGGATCT	CTGACATGAT	CCTGGGAGTCTTACTGCTCA	180	
182	tggcgtgttggt	attgtaga	agacaaatggat	acagagccttgat	ggataaaagctctc	241	
181	TCGGCTCTTG	TGTAATTTAG	AAGCAAGAAATG	GATACAGAGCCCTT	GATCGATAAAAGTCTTC	240	
242	atgttggcactcaat	gtgcctta	caagaagatg	gccacaagaaggtt	gtgatcatcggg	301	
241	ATGTGGCACTCA	ATGTGCCCTT	CAACAAGAGATG	CCCAAGAGGGTT	TGTGATCATCGGG	300	
302	acagcaaaagtgt	ctctca	agagaaaaactgt	gaacctgt	ggttcccaatgctccacctg	361	
301	ACAGCAAAAGT	GTCTCTTCA	AGAGAAAAC	TGTGAAC	TGTGGTTC	CCCAATGCTCCACCTG	360
362	cttatgagaaact	ctctcg	agacaacgt	caccacacacctt	attcacctta	agagccagcga	421
361	CTTATGAGAAACT	CTCTCG	AACACAGT	CACCACCACTT	ATTCACCTT	AAGAGCCACGCA	420
422	gacacctggagacat	gctgaa	attatctctc	tccacactttt	gcttgaatttaata	acaagg	481
421	GACACCTG-AGAC	TGCTGAAATTTAT	TTTCTC-TCACAC	TCTTTGGT	TTGAATTTAAT	TACAGA	478
482	ccatctaatggtt	ctcccttggg	aatgggtgt	aggnaaaaat	gncacccactct	aatna	541
479	C-ATCTAATG-TC	TCTCC-TTTGG	-AATGG-TGT	AGAAAAAT	TGCAAGCCCAT	CTCTAATAA	533
542	taaggtcagtggt	gataaaatttt	tgtagg	ccgact		576	
534	TAAG-TCAGTG-T	TAAAAATTTAG	-TAGGTC	CGCTG		565	

RESULT	5	standard; RNA; EST; 522 BP.
AC	H5510755	
AC	H975107	
OT	14-DEC-1995	(Rel. 46, Created)
OT	14-DEC-1995	(Rel. 46, Last updated, Version 1)
DE	yll1h07.s1	Homo sapiens cDNA clone 251965 3' similar to contains
DE	Alu	repetitive element;.
EST.		
OS	Homo sapiens	(human)
OC	Eukaryota; Animalia;	Metazoa; Chordata; Vertebrata; Mammalia;
OC	Theria; Eutheria; Primates;	Haplorhini; Catarrhini; Homiidae.
RP	11	
RP	1-522	
RA	Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,	
RA	Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,	
RA	Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,	
RA	Trevaskis E., Waterston R., Williamson A., Wohldmann P., Wilson R.,	
AT	"The WashU-Merck EST Project";	
AT	Unpublished.	
UL	Contract: Wilson RK WashU-Merck EST Project Washington University	
CC	School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,	
CC	MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:	
CC	est@wustl.wustl.edu High quality sequence stops: 313 Source: IMAGE	
CC	Consortium, LNL This clone is available royalty-free through LNL	
CC	; contact the IMAGE Consortium (info@image.llnl.gov) for further	
CC	information. NCBI gi: 1118395	
PH	Key	Location/Qualifiers
FT	source	1..522

```

FT      /organism="Homo sapiens"
FT      /clone="251965"
FT      /note="human"
FT      <1..>522
SQ      Sequence 522 BP; 143 A; 109 C; 128 G; 135 T; 7 other;

Query Match      30.9%; Score 482; DB 134; Length 522;
Best Local Similarity 97.7%; Pred. No. 0.00e+00;
Matches 507; Conservative 0; Mismatches 8; Indels 4; Gaps 4;

Db      4 gaatctctgatcctttatgattcttctgcacattctgtgagagggatttgggaacca 63
Cp      1522 GAATCTCTGATCCTTTATGATTTCTTCTGCACATTTCTGAGAGGATTTGGGAACCA 1463

Db      64 agagagctaaattgccataggtacatttaaggcagtcacacaggcatcagcactaaggcag 123
Cp      1462 AGAGAGCTAAATTGCCATAGGTACATTTAAGGCAGTACACAGGCTACACACTTAAGGCAG 1403

Db      124 tacatagtagctttactattttaatcctctaccatttctgtgttagtagaaaaatc 183
Cp      1402 TACATAGTAGCTTTACTATTATTAATCTCTTCTACACATTTCTGAGAGGATTTGGGAACCA 1343

Db      184 ttgtcattagtagaattgagccttgaagtgaagtgaattagcgtgattat 243
Cp      1342 TTGTCATTAGTTAGATAGCATTTGAGCCTTGAAAGTGAATGTAAATTAGAGCTGGGTATAT 1283

Db      244 gtcttacctaaagatatagatccagccaggcgtgggtcctcagcctgtaatccagcac 303
Cp      1282 GTCTTACCTAAGATATAGGATCCAGCCAGCGTGTGGCTCAGCCCTGTAATTCAGCAC 1223

Db      304 tttggaggctgagcggcgagatcaactgagtcaggagtttgagaccagcctgacca 363
Cp      1222 TTT-GGAGGCTGAGCGGGCAGATCAGCTGAGTCAGGAGTTTGAGACCGCTGACCA 1164

Db      364 acatggngaaacccgcgtctactaaaactacaaattactcagcanagtggcgacn 423
Cp      1163 ACATGG-AGAAACCCCGTCTCTACTTAAACTACAAAAATTAGTCAGGCTAGTGGCGACG 1105

Db      424 c-tgaatccagctnctcaggagctgaaggcaggagaattgcttaacnnggagcgg 482
Cp      1104 CCGTAATCCGCTACTACTAGAGGCTAAGCGAGGAATTGCTTGAACCTGGGAGCGG 1045

Db      483 aggttatggngagccagatcgcggccattgcactccag 521
Cp      1044 AGTTATGGTGACCCACATCGCG-CCATTGCACTCCAG 1007
```

```

RESULT 6
LOCUS      H97510      522 bp      mRNA      EST      12-DEC-1995
DEFINITION      yw11h07.s1 Homo sapiens cDNA clone 251965 3' similar to contains
                Alu repetitive element;.
ACCESSION      H97510
NID            g1118395
KEYWORDS      EST.
SOURCE      human clone=251965 primer=m13 -40 forward library=Soares melanocyte
                2NDHM vector-pT73D (Pharmacia) with a modified polylinker
                host=DH10B (ampicillin resistant) Reitel=Not I Rsite2=Eco RI Male.
                1st strand cDNA was primed with a Not I - oligo(dT) primer
                [5'-TGTACCAATCTGAAGTCGAGCGCGCCGAGTTTTTTTTTTTTTTT-3'],
                double-stranded cDNA was size selected, ligated to Eco RI adapters
                (Pharmacia), digested with Not I and cloned into the Not I and Eco
                RI sites of a modified pT73 vector (Pharmacia). Library
                constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal
                foreskin melanocytes (F5374) was kindly provided by Dr. Anthony P.
```

```

ORGANISM      Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 522)
AUTHORS      Hillier,L., Clark,N., Dubuque,T., Ellieton,K., Hawkins,M.,
                Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
                Parsons,J., Rifkin,L., Rohlffing,T., Soares,M., Tan,F.,
                Trevaakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
                Wilson,R.
TITLE      The WashU-Merck EST Project
JOURNAL      Unpublished (1995)
COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 313
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
```

```

NCBI gi: 1118395      Location/Qualifiers
source      1..522
              /organism="Homo sapiens"
              /clone="251965"
              /note="human"
              <1..>522
mRNA
BASE COUNT      143 a 109 c 128 g 135 t 7 other
ORIGIN

Query Match      30.9%; Score 482; DB 111; Length 522;
Best Local Similarity 97.7%; Pred. No. 0.00e+00;
Matches 507; Conservative 0; Mismatches 8; Indels 4; Gaps 4;

Db      4 gaatctctgatcctttatgattcttctgcacattctgtgagagggatttgggaacca 63
Cp      1522 GAATCTCTGATCCTTTATGATTTCTTCTGCACATTTCTGAGAGGATTTGGGAACCA 1463

Db      64 agagagctaaattgccataggtacatttaaggcagtcacacaggcatcagcactaaggcag 123
Cp      1462 AGAGAGCTAAATTGCCATAGGTACATTTAAGGCAGTACACAGGCTACACACTTAAGGCAG 1403

Db      124 tacatagtagctttactattttaatcctctaccatttctgtgttagtagaaaaatc 183
Cp      1402 TACATAGTAGCTTTTACTATTATTAATCTCTTCTACACATTTCTGAGAGGATTTGGGAACCA 1343

Db      184 ttgtcattagtagaattgagccttgaagtgaagtgaattagcgtgattat 243
Cp      1342 TTGTCATTAGTTAGATAGCATTTGAGCCTTGAAAGTGAATGTAAATTAGAGCTGGGTATAT 1283

Db      244 gtcttacctaaagatatagatccagccaggcgtgggtcctcagcctgtaatccagcac 303
Cp      1282 GTCTTACCTAAGATATAGGATCCAGCCAGCGTGTGGCTCAGCCCTGTAATTCAGCAC 1223

Db      304 tttggaggctgagcggcgagatcaactgagtcaggagtttgagaccagcctgacca 363
Cp      1222 TTT-GGAGGCTGAGCGGGCAGATCAGCTGAGTCAGGAGTTTGAGACCGCTGACCA 1164
```

Db 364 acatgnggaacccnctactactaaactacaaatagtcaggcanagtgccgacn 423  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Cp 1163 ACATGG-AGAACCCCGTCTCTACTTAAACTACAAAATAGTCAGGCATAGTGGCGCAG 1105  
Db 424 c-tgtaatcccaagctnctcaggaggtgaagcaggagaattgcttgaacnnggagcg 482  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Cp 1104 CCTGAATCCCACTACTACGAGGCTAAGCGAGGAGAAATTCCTTGACCTGGAGGCGG 1045  
Db 483 aggttatgmgagccagatcgccgcatgcactccag 521  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Cp 1044 AGGTTATGCTGACCAAGATCGCG-CCATTGCACCTCCAG 1007

## RESULT 7

LOCUS H97498 540 bp mRNA EST 12-DEC-1995  
DEFINITION yw11f07.s1 Homo sapiens cDNA clone 251941 3' similar to contains  
Alu repetitive element;

ACCESSION H97498

NID g1118383

KEYWORDS EST.

SOURCE human clone=251941 primer=ml3 -40 forward library=Soares melanocyte  
2NHM vector=pT73D (Pharmacia) with a modified polylinker  
host=DH10B (ampicillin resistant) Reitel=Not I Reite2=Eco RI Male.  
1st strand cDNA was primed with a Not I - oligo (dT) primer  
[5'-TCTTACCACTGAGTGGAGCGCGCGAGTTTTTTTTTTTTTTT-3'],  
double-stranded cDNA was size selected, ligated to Eco RI adapters  
(Pharmacia), digested with Not I and cloned into the Not I and Eco  
RI sites of a modified pT73 vector (Pharmacia). Library  
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal  
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.  
Albino.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;  
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
Sarcopterygia; Chonata; Tetrapoda; Amniota; Mammalia; Theria;  
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 540)  
AUTHORS  
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,  
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,  
Trevasaki,E., Waterston,R., Williamson,A., Wohlmann,P. and  
Wilson,R.

TITLE The WashU-Merck EST Project

JOURNAL Unpublished (1995)

## COMMENT

Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
High quality sequence stops: 377  
Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.

NCBI gi: 1118383

Location/Qualifiers

1..540

/organism="Homo sapiens"

/clone="251941"

/note="human"

## FEATURES

source

mRNA <1..>540  
BASE COUNT 153 a 111 c 135 g 136 t 5 others  
ORIGIN  
Query Match 30.0%; Score 468; DB 111; Length 540;  
Best Local Similarity 96.8%; Pred. No. 0.00e+00;  
Matches 520; Conservative 0; Mismatches 9; Indels 8; Gaps 7;  
Db 3 gaatctgatctcttattgtcttgcacattctgtgagaggatttgggaacca 62  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Cp 1522 GAATCTCTGATTCCTTTATGATTTCTTCTGCACATTTCTTGTGAGAGGATTTGGGAACCA 1463  
Db 63 agagagctaaattgcataggtacatttaaggcagtcacacaggcatcagcactaaggcag 122  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Cp 1462 AGAGAGCTAAATTTGCCATAGGTACATTTTAAAGGCAGTACACAGGCATCAGCCTAAGGCAG 1403  
Db 123 tacatagtagctttactatttaaatcctctaccattctcggtttagtagaaaaaac 182  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Cp 1402 TACATAGTAGCTTTTACTTTATTTAAATCCTTCTTACCAATTTCTGCTTTAGTAGAAAATAC 1343  
Db 183 ttgtcattagtagaataagcattgagcctgaagtgaatgaattagactgcgttatan 242  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Cp 1342 TTGCTATTAGTATAGATAGCATTGAGCCCTTGAGTGAATGTAATTAATAGACTGCGTTATAT 1283  
Db 243 gtcttaccatgaatagatccagccaggcgtggtgcacgcctgaattccagcac 302  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Cp 1282 GTCTTACCTAAGATATAGGATCCAGCCAGGCGTGGTCCAGCCCTGTAATTCAGCAC 1223  
Db 303 ttggaggagctgagcggcgagatcacctgaggtcagantttgagaccagcctggacc 362  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Cp 1222 TTTGGGAGGCTGAGGCGGCGAGATCACCTG-AGGTCAGGAGTTTGGAGCCAGCCTG-ACC 1165  
Db 363 aacatggagaaccccgctctctactaaactacaaaattagtcaggcatagtgccgacag 422  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Cp 1164 AACATGGAGAAACCCCGCTCTACTTAAAACTACAAAATTAAGTCAGGCATAGTGGCGCAGG 1105  
Db 423 c-tgtatcccaagctactcaggaggctgaagcaggagaattgcttgaacctggggaagcg 481  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Cp 1104 CCTGTAATCCAGCTACTTCAGGAGGCTAAGCGAGGAGTAATGCTTGACCTGCGG-AGCG 1047  
Db 482 cgaaggttaagggtgagccagatcgccantgcacttcacagctcgggnaaaaaa 538  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Cp 1046 -GGAGGTTATGG-TGAGCCAAAGATCGGCCATTGCACT-CCAGCCTGGGCAACAAA 993

## RESULT 8

ID H5498253 standard; RNA; EST; 540 BP.  
AC H97498;  
DT 14-DEC-1995 (Rel. 46, Created)  
DT 14-DEC-1995 (Rel. 46, Last updated, Version 1)  
DE yw11f07.s1 Homo sapiens cDNA clone 251941 3' similar to contains  
DE Alu repetitive element;  
KW EST.  
OS Homo sapiens (human)  
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;  
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.  
RN [1]  
RP 1-540  
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,  
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,  
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,  
RA Trevasaki E., Waterston R., Williamson A., Wohlmann P., Wilson R.;  
RT "The WashU-Merck EST Project";  
RL Unpublished.

2NHM vector-p773D (Pharmacia) with a modified polylinker host-DH10B (ampicillin resistant) RseI=Not I RseI2=Eco RI Male. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTACCAATCTGAAGTGGGAGCGCCGACGTTTTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p773 vector (Pharmacia). Library constructed by Bento Soares and M.Patima Bonaldo. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino.

**ORGANISM**  
Homo sapiens  
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonacta; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homidae; Homo.

**REFERENCE**  
1 (bases 1 to 473)

**AUTHORS**  
Hallier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevisan,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.

**TITLE**  
The WashU-Merck EST Project

**JOURNAL**  
Unpublished (1995)

**COMMENT**  
Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
High quality sequence stops: 363  
Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

**NCBI gi:** 1124758

**FEATURES**  
source  
1..473  
/organism="Homo sapiens"  
/clone="263806"  
/note="human"  
<...>473

**mRNA**  
127 a 97 c 97 g 151 t 1 others

**BASE COUNT**  
127 a 97 c 97 g 151 t 1 others

**ORIGIN**  
Query Match 28.3%; Score 441; DB 112; Length 473;  
Best Local Similarity 98.7%; Pred.No. 0.00e+00;  
Matches 467; Conservative 0; Mismatches 2; Indels 4; Gaps 4;  
1 catgcttgcaacatttattgatggagttttcccaatttaatttcatttcatttcctc 60  
|||||  
642 CATGCTTGCAACATTATTGATGGAGTTTCCCAATTTAAATTTTCATCATTTCTC 583  
|||||  
61 acatgattagtagctgtagcggaactactaaaattttaaacctgacttatattagagat 120  
|||||  
582 ACATGATTAGTAGCTGAGCGGAGCTACTAAAATTTTAACTGACTTATTATTAGAGAT 523  
|||||  
121 ggcttgatttttctacaccattccaaggaggaacattagatgtctgtattataatcaa 180  
|||||  
522 GGCTTGCAATTTTCTTACACCATTCGAAAGAGAACATTAGATGTCGTATTAATTCGA 463  
|||||  
181 gcaaaagtggagaaataatttcagcatgtctcaggtgtctcgctggctcttaaggtg 240  
|||||

Cp 462 GAAAAGTGTGAGAGAAATAATTTCAGCATGCTCAGGTGCTCGCTGGCTCTTAAGGTG 403  
 Db 241 gataagtggtggtgactgttctgagagagtttctcataagcaggtgagcattgggaa 300  
 Cp 402 AATAAGGTGGTGGTCTGACTCTTCAGAGAGTTCCTCATTAAGCAGGTGGAGCATGGGAA 343  
 Db 301 ccacaggtcacagttttctcttggaagagacactttgctgccgatgatcaaacctt 360  
 Cp 342 CCACAGGTTACAGTTTTCTCTTGAAGAGACACTTGTGCTGCCGATGATCAAAACCTT 283  
 Db 361 ctgtgggcattcttctgtaaggccacattngagtgccaacacacgcggtgagcag 473  
 Cp 282 CTGTGGGCATCTCTTCTTGAAGG-CACATT-GAGTGGCAC-ATG-AGACTTTTATCC 227  
 Db 421 atcaaggtctgtatccattctcttacaataccaacacgcggtgagcag 473  
 Cp 226 ATCAAGGCTGTATCCATTCTCTTCAATAACCAACAGCGATGAGCAG 174

## RESULT 10

ID HS091255 standard; RNA; EST; 473 BP.  
 AC N20091;  
 DT 17-DEC-1995 (Rel. 46, Created)  
 DT 17-DEC-1995 (Rel. 46, Last updated, Version 1)  
 DE yx36a12.s1 Homo sapiens cDNA clone 263806 3'.  
 KW EST.  
 OS Homo sapiens (human)  
 OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;  
 OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.  
 RN [1]  
 RP 1-473  
 RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,  
 RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,  
 RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,  
 RA Trevasakis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;  
 RT "The WashU-Merck EST Project";  
 RL Unpublished.  
 CC Contact: Wilson RK WashU-Merck EST Project Washington University  
 CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,  
 CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:  
 CC est@wustl.wustl.edu High quality sequence stops: 363 Source: IMAGE  
 CC Consortium, LBNL This clone is available royalty-free through LBNL  
 CC ; contact the IMAGE Consortium (info@image.llnl.gov) for further  
 CC information. NCBI gi: 1124758  
 FH Key Location/Qualifiers  
 FT source 1..473  
 FT /organism="Homo sapiens"  
 FT /clone="263806"  
 FT /note="human"  
 FT mRNA <1..>473  
 SQ Sequence 473 BP; 127 A; 97 C; 97 G; 151 T; 1 other;

Query Match 28.3%; Score 441; DB 129; Length 473;  
 Best Local Similarity 98.7%; Pred. No. 0.00e+00;  
 Matches 467; Conservative 0; Mismatches 2; Indels 4; Gaps 4;

Db 1 catgcatgcaacatttatgtaggagtttcccaatttaatttctcatctctc 60  
 Cp 642 CATGATGCAACATTTATGATGAGTTTCCCAATTATATTCTCATCATTTCTC 583  
 Db 61 acatgattagctgctagcgagacctactaaaatttaacactgacttattattagat 120  
 Cp 582 ACATGATTAGTACTGCTAGCGGAGCTACTATAAATTTTAACACTGACTTATTATTAGAT 523

Db 121 ggcctgcatctttctcacacattccaagagagacattatgatgtctgtattaaattcaa 180  
 Cp 522 GCTTGCAATTTTCTCACCATTCACAGGAGACATTAGATGCTGTATTAATTCAA 463  
 Db 181 gcaaaagtgtgagagaaataatttcagcatgtctcaggtgtctcgtcgctcttaaggtg 240  
 Cp 462 GCAAAAGTCTCAGAGAAATAATTTCAGCATGCTCAGGTGCTCGCTGGCTCTTAAGGTG 403  
 Db 241 gataagtggtggtgactgttctgagagagtttctcataagcaggtgagcattgggaa 300  
 Cp 402 AATAAGGTGGTGGTCTGACTCTTCAGAGAGTTCCTCATTAAGCAGGTGGAGCATGGGAA 343  
 Db 301 ccacaggtcacagttttcttctgaagagacactttgctgccgatgatcaaacctt 360  
 Cp 342 CCACAGGTTACAGTTTTCTCTTGAAGAGACACTTGTGCTGCCGATGATCAAAACCTT 283  
 Db 361 ctgtgggcattcttctgtaaggccacattngagtgccaacacacgcggtgagcag 473  
 Cp 282 CTGTGGGCATCTCTTCTTGAAGG-CACATT-GAGTGGCAC-ATG-AGACTTTTATCC 227  
 Db 421 atcaaggtctgtatccattctcttacaataccaacacgcggtgagcag 473  
 Cp 226 ATCAAGGCTGTATCCATTCTCTTCAATAACCAACAGCGATGAGCAG 174

## RESULT 11

LOCUS N24677 433 bp mRNA EST 28-DEC-1995  
 DEFINITION yx91a09.s1 Homo sapiens cDNA clone 269080 3'.  
 ACCESSION N24677  
 NID g1138827  
 KEYWORDS EST.  
 SOURCE human clone=269080 primer=ml3-40 forward library=Soares melanocyte  
 2Nblm vector=pT7T3D (Pharmacia) with a modified polylinker  
 host=DH10B (ampicillin resistant) Reitel=Not I Reite2=Eco RI Male.  
 1st strand cDNA was primed with a Not I - oligo (dT) primer  
 [5'-TGTACCAATCTGAAGTGGGCGCGCAGTGTGTGTGTGTGTGTGTGTGTGTGT-3'],  
 double-stranded cDNA was size selected, ligated to Eco RI adapters  
 (Pharmacia), digested with Not I and cloned into the Not I and Eco  
 RI sites of a modified pT73 vector (Pharmacia). Library  
 constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal  
 foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.  
 Albino.

## ORGANISM

Homo sapiens  
 Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;  
 Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
 Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;  
 Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.

## REFERENCE

1 (bases 1 to 433)

## AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
 Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and  
 Wilson, R.

## TITLE

The WashU-Merck EST Project

## JOURNAL

Unpublished (1995)

## COMMENT

Contact: Wilson RK  
 WashU-Merck EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.wustl.edu

High quality sequence stops: 315  
Source: IMAGE Consortium, L1NL  
This clone is available royalty-free through L1NL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 1138827 Location/Qualifiers  
source 1..433 /organism="Homo sapiens"  
/clone="269080"  
/note="human"  
mRNA <1..>433  
BASE COUNT 119 a 85 c 86 g 142 t 1 others  
ORIGIN

Query Match 27.1%; Score 423; DB 113; Length 433;  
Best Local Similarity 99.3%; Pred. No. 0.00e+00;  
Matches 430; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Db 1 atcatncaatgcaacattattgatggagtttcccaatttaatttctcatcattcc 60  
|||||  
Cp 644 ATCATGCAATGCAACATTTATGATGGAGTTTCCCAATTAATATTTCTCATTTCC 585  
|||||  
Db 61 tcacatgattagtagtctagcggaactactaaaattttaacactgaactattattagag 120  
|||||  
Cp 584 TCACATGATTAGTAGTCTAGCGGACCTACTAAAAATTTTAACACTGACTTATTATTAGAG 525  
|||||  
Db 121 atggttcatttttctcaacattccaaaggagaaacattagatgtctgtattaaattc 180  
|||||  
Cp 524 ATGGCTTCATTTTCTTACACCATTTCCAAAGGAGAACATTAGATGCTGTATTAAATTC 465  
|||||  
Db 181 aagcaaaagtgtgagagaataatttcagcatgtctcaggtgtctcgctggtcttaagg 240  
|||||  
Cp 464 AAGCAAAAGTGTGAGAGAAATATTTACAGATGCTTCAGGTGCTCGCTGGCTCTTAAGG 405  
|||||  
Db 241 tgaataaagggtggtgactgttctcgagagagtttctcataaagcaggtggagcattggg 300  
|||||  
Cp 404 TGAATAAGGTGGTGGTCACTGTTCTCCAGAGAGTTCCTATAAGCAGGTGGAGCATTTGGG 345  
|||||  
Db 301 aacacaggttcaacagtttttctcttggaagagacatttgcgtcccgatgatcaaac 360  
|||||  
Cp 344 AACACAGGTTCAACAGTTTCTTCTTGAAGGACACTTTTGCTGTCCCGAT-GATCAAAAC 286  
|||||  
Db 361 cctcttggtggcatcttctgttaaggcacattgagtgcacacatgaagactttttatcca 420  
|||||  
Cp 285 CTTCCTTGTGGGATCTTCTTCTTGAAGGACACTTTTGCTGTCCCGAT-GATCAAAAC 226  
|||||  
Db 421 tcaaggtctctgta 433  
|||||  
Cp 225 TCAGGCTCTGTGA 213  
|||||

RESULT 12  
ID HS677260 standard; RNA; EST; 433 BP.  
AC N24677;  
DT 30-DEC-1995 (Rel. 4.6, Created)  
DT 30-DEC-1995 (Rel. 4.6, Last updated, Version 1)  
DE yx91a09.s1 Homo sapiens cDNA clone 269080 3'.  
KW EST.  
OS Homo sapiens (human)  
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;  
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.  
RN [1]  
RP 1-433

RA Hillier L., Clark N., Duboue T., Elliston K., Hawkins M.,  
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,  
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,  
RA Trevaekis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;  
RT "The WashU-Merck EST Project";  
RL Unpublished.

CC Contact: Wilson RK WashU-Merck EST Project Washington University  
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,  
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:  
CC est@wustl.wustl.edu High quality sequence stops: 315 Source: IMAGE  
CC Consortium, L1NL This clone is available royalty-free through L1NL  
CC ; contact the IMAGE Consortium (info@image.llnl.gov) for further  
CC information. NCBI gi: 1138827

FH Key Location/Qualifiers  
FH source 1..433 /organism="Homo sapiens"  
/clone="269080"  
/note="human"  
mRNA <1..>433  
SQ Sequence 433 BP; 119 A; 85 C; 86 G; 142 T; 1 other;

Query Match 27.1%; Score 423; DB 135; Length 433;  
Best Local Similarity 99.3%; Pred. No. 0.00e+00;  
Matches 430; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Db 1 atcatncaatgcaacattattgatggagtttcccaatttaatttctcatcattcc 60  
|||||  
Cp 644 ATCATGCAATGCAACATTTATGATGGAGTTTCCCAATTAATATTTCTCATTTCC 585  
|||||  
Db 61 tcacatgattagtagtctagcggaactactaaaattttaacactgaactattattagag 120  
|||||  
Cp 584 TCACATGATTAGTAGTCTAGCGGACCTACTAAAAATTTTAACACTGACTTATTATTAGAG 525  
|||||  
Db 121 atggttcatttttctcaacattccaaaggagaaacattagatgtctgtattaaattc 180  
|||||  
Cp 524 ATGGCTTCATTTTCTTACACCATTTCCAAAGGAGAACATTAGATGCTGTATTAAATTC 465  
|||||  
Db 181 aagcaaaagtgtgagagaataatttcagcatgtctcaggtgtctcgctggtcttaagg 240  
|||||  
Cp 464 AAGCAAAAGTGTGAGAGAAATATTTACAGATGCTTCAGGTGCTCGCTGGCTCTTAAGG 405  
|||||  
Db 241 tgaataaagggtggtgactgttctcgagagagtttctcataaagcaggtggagcattggg 300  
|||||  
Cp 404 TGAATAAGGTGGTGGTCACTGTTCTCCAGAGAGTTCCTATAAGCAGGTGGAGCATTTGGG 345  
|||||  
Db 301 aacacaggttcaacagtttttctcttggaagagacatttgcgtcccgatgatcaaac 360  
|||||  
Cp 344 AACACAGGTTCAACAGTTTCTTCTTGAAGGACACTTTTGCTGTCCCGAT-GATCAAAAC 286  
|||||  
Db 361 cctcttggtggcatcttctgttaaggcacattgagtgcacacatgaagactttttatcca 420  
|||||  
Cp 285 CTTCCTTGTGGGATCTTCTTCTTGAAGGACACTTTTGCTGTCCCGAT-GATCAAAAC 226  
|||||  
Db 421 tcaaggtctctgta 433  
|||||  
Cp 225 TCAGGCTCTGTGA 213  
|||||

RESULT 13  
ID HS796259 standard; RNA; EST; 396 BP.  
AC N24796;  
DT 30-DEC-1995 (Rel. 4.6, Created)  
DT 30-DEC-1995 (Rel. 4.6, Last updated, Version 1)

DE yx97f10.s1 Homo sapiens cDNA clone 269707 3'.

KW EST.

OS Homo sapiens (human)

OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;

OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.

RN [1]

RP 1-396

RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,

RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,

RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,

RA Trevasakis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;

RT "The WashU-Merck EST Project";

RL Unpublished.

CC Contact: Wilson RK WashU-Merck EST Project Washington University

CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,

CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:

CC est@wustl.edu High quality sequence stops: 310 Source: IMAGE

CC Consortium, L1NL This clone is available royalty-free through L1NL

CC ; contact the IMAGE Consortium (info@image.llnl.gov) for further

CC information. NCBI gi: 1138946

FH Key Location/Qualifiers

FT source 1..396 /organism="Homo sapiens"

FT /clone="269707"

FT /note="human"

FT mRNA <1..>396

FT Sequence 396 BP; 108 A; 75 C; 80 G; 131 T; 2 other;

Query Match 25.3%; Score 394; DB 136; Length 396;

Best Local Similarity 99.7%; Pred. No. 0.00e+00;

Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 atcatnctgcaacatttattgatggagttttcccaatttaatttctcatcttcc 60

|||||

Cp 644 ATCATGCAATTCGCAACATTTATGAGGAGTTTCCCAATTTAATATTTCTCATATTC 585

Db 61 tcacatgattagctgctagcggacctactaaattttaacactgacttatttagag 120

|||||

Cp 584 TCACATGATTAGTACTGCTAGCGGACCTACTAAATTTTAACTGACTTATTATTAGAG 525

Db 121 atggcttgcatcttctacacattccaaaggagacattagatgtctgtattaatcc 180

|||||

Cp 524 ATGGCTTCATTTTCTTACACCATTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 465

Db 181 aagcaaaagtgtgagagaataatttcacagctgtctcaggtgtctcgctggctctaaag 240

|||||

Cp 464 AGCCTAAGTGTGAGAGAAATATTTACAGTGTCTCAGGTGTCTCGCTGGCTCTTAAG 405

Db 241 tgaataaagtggtggtgactgttctgcagagagtttctcataagcaggtgagcattggg 300

|||||

Cp 404 TGAATAAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 345

Db 301 aaccacaggttcaaggtttttcttcttgaagacacatttctgtccgatgataaacc 360

|||||

Cp 344 AACCAAGGTTACAGATTTTCTTCTTGAAGAGACACTTTGCTGCTCCGATGATCAACCC 285

Db 361 ttcttggtggcactcttcttgaagacattgag 395

|||||

Cp 284 TTCTTGTGGGATCTTCTTGTGTAAGGCACATTGAG 250

RESULT 14

LOCUS N24796 396 bp mRNA EST 28-DEC-1995

DEFINITION yx97f10.s1 Homo sapiens cDNA clone 269707 3'.

ACCESSION N24796

NID g1138946

KEYWORDS EST.

SOURCE human clone=269707 primer=ml3 -40 forward library=Soares melanocyte

2N8M vector=PT73D (Pharmacia) with a modified polylinker

host=DH10B (ampicillin resistant) Reitel=Not I Reite2=Eco RI Male.

1st strand cDNA was primed with a Not I - oligo(dT) primer

[5'-TGTACCAATCTCAAGTGGAGCGCGCCAGTTTTTTTTTTTTTTT-3'],

double-stranded cDNA was size selected, ligated to Eco RI adapters

(Pharmacia), digested with Not I and cloned into the Not I and Eco

RI sites of a modified PT73 vector (Pharmacia). Library

constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal

foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.

Albino.

ORGANISM Homo sapiens

Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;

Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;

Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;

Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 396)

AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,

Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,

Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,

Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P., and

Wilson, R.

TITLE The WashU-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT

Contact: Wilson RK

WashU-Merck EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

High quality sequence stops: 310

Source: IMAGE Consortium, L1NL

This clone is available royalty-free through L1NL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 1138946

Location/Qualifiers

1..396

/organism="Homo sapiens"

/clone="269707"

/note="human"

mRNA <1..>396

BASE COUNT 108 a 75 c 80 g 131 t 2 others

ORIGIN

Query Match 25.3%; Score 394; DB 113; Length 396;

Best Local Similarity 99.7%; Pred. No. 0.00e+00;

Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 atcatnctgcaacatttattgatggagttttcccaatttaatttctcatcttcc 60

|||||

Cp 644 ATCATGCAATTCGCAACATTTATGAGGAGTTTCCCAATTTAATATTTCTCATATTC 585

Db 61 tcacatgattagctgctagcggacctactaaattttaacactgacttatttagag 120

|||||

Cp 584 TCACATGATTAGTACTGCTAGCGGACCTACTAAATTTTAACTGACTTATTATTAGAG 525

Db 121 atggcttgcatcttctacacattccaaaggagacattagatgtctgtattaatcc 180



Cp 524 ATGGCTTGCATTTTCTACACCATTCCAAGAGAGAACATTAGATGCTGTATTAAATTC 465  
Db 181 aagcaaaagtgtgagagaataatttcagcatgtctcaggtgtctcgctggtcttaagg 240  
Cp 464 AAGCAAAAGTGTGAGAGAAATAATTCAGCATGCTCAGGTGTCTCGCTGGCTCTTAAGG 405  
Db 241 tgaataaagtggtggtgactgttcgcagagagatttctcataagcagggtgagcattggg 300  
Cp 404 TGAATAAGGTGGTGTGACTGCTCTGCAGACAGTTCCTATAAGCAGGTGGAGCAATTGGG 345  
Db 301 aaccacaggttcacagttttcttctgaagagacacttctgtcccgatgatcaaaccc 360  
Cp 344 AACCCAGGTTACAGTTTCTTCTTGAAGAGACACTTTGCTGTCCGATGATCAAAACC 285  
Db 361 ttctgtgggcattcttcttgaaggcacattgag 395  
Cp 284 TTCTTGTGGCATCTCTTGTGTAAGGCACATTGAG 250

## RESULT 15

LOCUS N26562 401 bp mRNA EST 29-DEC-1995  
DEFINITION Yx62h05.s1 Homo sapiens cDNA clone 266361 3'.  
ACCESSION N26562  
NTD q1140910  
KEYWORDS EST.  
SOURCE human clone=266361 primer=ml3 -40 forward library=Soares melanocyte  
2NBHM vector=pT7T30 (Pharmacia) with a modified polylinker  
host=DH10B (ampicillin resistant) Reitel=Not I Rsite2=Eco RI Male.  
1st strand cDNA was primed with a Not I - oligo(dT) primer  
[5'-TGTACCAATCTCAAGTGGAGCGCGCGAGTTTTTTTTTTTTTTT-3'],  
double-stranded cDNA was size selected, ligated to Eco RI adapters  
(Pharmacia), digested with Not I and cloned into the Not I and Eco  
RI sites of a modified pT7T3 vector (Pharmacia). Library  
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal  
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.  
Albino.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;  
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;  
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 401)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
Parsons, J., Rifkin, L., Rohlfs, T., Soares, M., Tan, F.,  
Trevas, E., Waterston, R., Williamson, A., Wohlmann, P. and  
Wilson, R.

## TITLE

The WashU-Merck EST Project

## JOURNAL

Unpublished (1995)

## COMMENT

Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu  
High quality sequence stops: 305  
Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.

NCBI gi: 1140910

FEATURES  
source Location/Qualifiers  
1..401  
/organism="Homo sapiens"  
/clone="266361"  
/note="human"  
mRNA  
BASE COUNT 110 a 79 c 82 g 129 t 1 others  
ORIGIN  
Query Match 25.1%; Score 392; DB 114; Length 401;  
Best Local Similarity 99.5%; Pred. No. 0.00e+00;  
Matches 393; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Db 1 atcgtcattgcaaacatttattgattgaggtttcccaatttataatttctcattcc 60  
Cp 644 ATCATGCAATTCGAACATTTATTGATGGAGTTTCCCAATTTTAATATTTCTCATCATTC 585  
Db 61 tcacatgattagtagtctgtagcgacctactaaattttaacacactgaattattagag 120  
Cp 584 TCACATGATTAGTACTGCTACGCGACCTACTAAATTTTAACTGACTTATTATTAGAG 525  
Db 121 atggtcttgcatttttctcacaccattccaaaggagacattagatgtctgtattaaattc 180  
Cp 524 ATGGCTTGCATTTTCTTACACCATTTCCAAAGAGAGAACATTAGATGCTGTATTAAATTC 465  
Db 181 aagcaaaagtgtgagagaataatttcagcatgtctcaggtgtctcgctggccttaagg 240  
Cp 464 AAGCAAAAGTGTGAGAGAAATAATTTTACAGCATGCTCAGGTGTCTCGCTGGCTCTTAAGG 405  
Db 241 tgaataaagtggtggtgactgttcgcagagagatttctcataagcagggtgagcattggg 300  
Cp 404 TGAATAAGGTGGTGTGACTGCTTCGACAGACAGTTTCTCATAAGCAGGTGGAGCAATTGGG 345  
Db 301 aaccacaggttcacagtttttctcttgaagagacacttctgtcccgatgatcaaaccc 360  
Cp 344 AACCCAGGTTACAGTTTCTTCTTGAAGAGACACTTTTGTGTCGCCGATGATCAAAACC 285  
Db 361 ttctgtgggcattcttctgttaaggcacattgag 395  
Cp 284 TTCTTGTGGCATCTCTTGTGTAAGGCACATTGAG 250

Search completed: Tue Aug 27 08:43:15 1996

Job time : 642 secs.



\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

(TM)

\*\*\*\*\*

MPsrch\_nm n.a. - n.a. database search, using Smith-Waterman algorithm  
Run on: Tue Aug 27 08:07:40 1996; MasPar time 1034.82 Seconds  
Tabular output not generated.  
1145.061 Million cell updates/sec

Title: >US-08-231-565A-1  
Description: (1-1559) from US08231565A.seq  
Perfect Score: 1559  
N.A. Sequence: 1 AGCAGACAGAGACTCTCAT.....AAAAAAAAAAAAAAAAAAAA 1559  
Comp: TCGCTCTCTCTCTGAGAGTA.....TTTTTTTTTTTTTTTTTTTT

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 270440 seqs, 380027776 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: embl-new3  
1:BCT 2:FUN 3:INV1 4:INV2 5:INV3 6:MAM 7:ORG 8:PLN  
9:PRI1 10:PRI2 11:PRO 12:ROD 13:SYN 14:UNC 15:VRT  
16:VIR1 17:VIR2

Database: genbank92  
18:BCT1 19:BCT2 20:BCT3 21:BCT4 22:BCT5 23:BCT6 24:BCT7  
25:BCT8 26:BCT9 27:INV1 28:INV2 29:INV3 30:INV4 31:INV5  
32:INV6 33:MAM1 34:MAM2 35:PAT1 36:PAT2 37:PAT3 38:PHG  
39:PLN1 40:PLN2 41:PLN3 42:PLN4 43:PLN5 44:PLN6 45:PLN7  
46:PLN8 47:PRI1 48:PRI2 49:PRI3 50:PRI4 51:PRI5 52:PRI6  
53:PRI7 54:PRI8 55:PRI9 56:PRI10 57:PRI11 58:PRI12  
59:PRI13 60:ROD1 61:ROD2 62:ROD3 63:ROD4 64:ROD5 65:ROD6  
66:ROD7 67:STR 68:SYN 69:UNA 70:VRL1 71:VRL2 72:VRL3  
73:VRL4 74:VRL5 75:VRL6 76:VRL7 77:VRT1 78:VRT2 79:VRT3

Database: genbank-new1  
80:BCT 81:INV1 82:INV2 83:MAM 84:PHG 85:PLN 86:PRI  
87:ROD 88:STR 89:SYN 90:UNA 91:VRL 92:VRT

Database: u-embl45\_92  
93:part1

Statistics: Mean 12.154; Variance 9.393; scale 1.294

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description	Pred. No.
1	1524	97.8	1524	52	HSU06452	Human melanoma antigen
2	645	41.4	752	52	HSU06654	Human differentiation
3	238	15.3	291	68	HSU14572	***ALU WARNING: Human
4	238	15.3	35962	9	HSBLGR1	Human ABL gene, exon
5	238	15.3	35962	49	HSABLR1	Human ABL gene, exon
6	238	15.3	35962	86	HSABLR1	Human ABL gene, exon
7	237	15.2	9457	49	HSATPCP1	H.sapiens gene for mi
8	232	14.9	6901	51	HSLCATG	H.sapiens gene for le
9	232	14.9	6905	52	HSTMP04	Human thymopoietin (t
10	232	14.9	13863	52	HSPROSCY	H.sapiens genes for p
11	232	14.9	16851	10	HSL69F7B	Human DNA sequence fr
12	233	14.9	219447	10	HSFLNG6PD	Homo sapiens chromoso
13	232	14.9	219447	10	HSFLNG6PD	Homo sapiens chromoso
14	230	14.8	1171	58	HUMTCTPV2	Human Pro-tRNA and Va
15	230	14.8	1846	10	HSL174G8C	Human DNA sequence fr
16	231	14.8	1929	51	HSL139H8	Human DNA sequence fr
17	229	14.7	291	68	HSU14573	***ALU WARNING: Human
18	229	14.7	36594	58	HUMTPA	Human tissue plasmino
19	227	14.6	34641	57	HUMRCC1	Human RCC1 gene, comp
20	226	14.5	3810	10	HSL174G8B	Human DNA sequence fr
21	226	14.5	4555	53	HSUBAS2G	Human UBA52 gene codi
22	226	14.5	18004	54	HUMATP5AS	Human gene for ATP sy
23	226	14.5	32351	56	HUMLHDC	Human gene for L-hist
24	226	14.5	38759	51	HSN13E1	Human DNA sequence fr
25	226	14.5	84539	49	HSABLR3	Human proto-oncogene
26	226	14.5	84539	9	HSBLGR3	Human proto-oncogene
27	226	14.5	84539	86	HSABLR3	Human proto-oncogene
28	224	14.4	28265	10	HSRJ14	Human DNA sequence fr
29	224	14.4	37695	86	HSL95E6	Human DNA sequence fr
30	224	14.4	37695	10	HSL95E6	Human DNA sequence fr
31	224	14.4	37785	58	HUMXRCC1G	Human XHCC1 DNA repl
32	224	14.4	38409	55	HUMDXS455A	Human cosmid insert c
33	225	14.4	40087	10	HSU141H12	Human DNA sequence fr
34	221	14.2	6870	56	HUMIL11A	Human interleukin 11
35	222	14.2	17538	10	HSMOGG	H.sapiens gene for my
36	222	14.2	17538	51	HSMOGG	H.sapiens gene for my
37	222	14.2	29103	10	HSL165D7	Human DNA sequence fr
38	222	14.2	29103	86	HSL165D7	Human DNA sequence fr
39	221	14.2	35473	86	HSL30G1	Human DNA sequence fr
40	222	14.2	40592	86	HSL191F1	Human DNA sequence fr
41	222	14.2	40592	10	HSL191F1	Human DNA sequence fr
42	222	14.2	47934	53	HSU24498	Human autosomal domin
43	222	14.2	53522	57	HUMPRD1GEN	Homo sapiens polycyst
44	221	14.2	152141	86	HSU07000	Human breakpoint clus
45	221	14.2	152141	9	HS07000	Human breakpoint clus

## ALIGNMENTS

RESULT 1  
LOCUS HSU06452 1524 bp mRNA PRI 25-JUN-1994  
DEFINITION Human melanoma antigen recognized by T-cells (MART-1) mRNA.  
ACCESSION U06452  
NID g476131  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;

Aug 27 08:17

US-08-231-565A-1.rge

3

REFERENCE 1 (bases 1 to 1524)  
AUTHORS Kawakami, Y., Elyahu, S., Delgado, C.H., Robbins, P.F., Rivoltini, L., Topalian, S.L., Miki, T. and Rosenberg, S.A.  
TITLE Cloning of the gene coding for a shared human melanoma antigen recognized by autologous T cells infiltrating into tumor  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91, 3515-3519 (1994)  
MEDLINE 94224770

REFERENCE 2 (bases 1 to 1524)  
AUTHORS Kawakami, Y.  
TITLE Direct Submission  
JOURNAL Submitted (08-FEB-1994) Yutaka Kawakami, Surgery Branch, National Cancer Institute, National Institutes of Health, 9000 Rockville Pike, Bldg. 10, Rm. 2B42, Bethesda, MD 20892, USA  
COMMENT NCB1 gi: 476131

FEATURES  
source Location/Qualifiers  
1..1524  
/organism="Homo sapiens"  
/cell\_line="melanoma cell line, 501mel"  
/tissue\_type="melanoma"  
54..410  
/gene="MART-1"  
/standard\_name="melanoma antigen recognized by T-cells"  
/note="NCBI gi: 476132"  
/codon\_start=1  
/db\_xref="PID:9476132"  
/translation="MPREDARFYGYPKKGHSYVTAEEAAGIGLTVILGVLLIG  
CWYCHRRNGYBALMDKSLHVGTCALTRPCQEGFDHRDVKVSLQKNCPEVVPNAPP  
AYELISAEQSPPPYSP"  
BASE COUNT 435 a 330 c 324 g 435 t  
ORIGIN

Query Match 97.8%; Score 1524; DB 52; Length 1524;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 1524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 acgacagaggactctcattaaggaggtgtcctgtgccctgacctacagaatgccaa 60  
|||||  
Qy 1 AGCAGACAGGACTCTCATTAAGGAAGGTCTCTGTGCCCTGCACCTACAAGATGCCAA 60  
Db 61 gagaagatgctcactctctatggttaaccccaagaaggcgccagccactcttaccca 120  
|||||  
Qy 61 GAGAAATGCTCATTTCATCTATGTTATGCCCAAGAAGGGGCGCGCCTTTACACCA 120  
Db 121 cggctgaagaggccgtgggatcggtcctgacagdgatcctggagcttactgetca 180  
|||||  
Qy 121 CGGCTGAAGAGGCCGCTGGGATCGGCATCTCTGACATGTCCTGGGATCTTACTGTCA 180  
Db 181 tcgctgtgtggtattgtgaagacgaatggatagagccctgtgataaaagtcttc 240  
|||||  
Qy 181 TCGGCTGTGTGATTGTGAAGACGAATGGATACAGACCCCTTCATGATATAAAGTCTTC 240  
Db 241 atgttggcaactcaatgtccttaacgaagatgcccacaagaagggtttgatcggg 300  
|||||  
Qy 241 ATTGTGGCACTCAATTGCGCTTAACAAGAATGCCCAAGAAGGGTTTGATCATCGGG 300  
Db 301 acagcaagtgctctctcaagagaaaaacttgaaacctgtggttcccaatgctccactg 360  
|||||  
Qy 301 ACAGCAAGTGCTCTCTCAAGAGAAAACTGTGAACCTGTGTTCCCAATGCTCCACCTG 360  
Db 361 cttatgagaactctctcgagaacagctcaccacacaccttattcaccttaagagcagcga 420  
|||||  
Qy 361 CTTATGAGAACTCTCTGACAGACAGTCAACACCACTTATTCACCTTAAGAGCCGCA 420

Aug 27 08:17

US-08-231-565A-1.rge

4

Db 421 gacacctgagacagtgcgaatattttctcaccacttttgcctgaatttaatacagaca 480  
|||||  
Qy 421 GACACCTGAGACAGTGTGAATATTTCCTCTCACACTTTTGTCTGAATTTAATACAGACA 480  
Db 481 tctaattgtctctcttgggaatggtgtaggaaaaatgaagccatctctaataaagtca 540  
|||||  
Qy 481 TCTAATGTTCTCTCTTGGAAATGTTGTAGGAAAAATCGAAGCCATCTCTAATAAATAGTCA 540  
Db 541 gtgttaaaatttttagtagtcgctagtcagctactaatcatgtgaggaatgatgagaat 600  
|||||  
Qy 541 GTGTAAAAATTTTAGTAGTCCGCTAGCAGTACTTAATCATGTGAGGAAATGATCAGAAAT 600  
Db 601 attaaatgggaaaaactccatcaataaattgttgcgaatgcacatgatactatctgtccagag 660  
|||||  
Qy 601 ATTAAATTGGGAAACTCCATCAATAAATGTTGCAATGCATGATCTATCTGTGCCAGAG 660  
Db 661 gtaatgttagtaaatccatggtgtattttctgagagacagaatcaagtggtgattctg 720  
|||||  
Qy 661 GTAATGTTAGTAATCCATGCTGTATTTCCTGAGAGACAGAAATTCAGTGGCTATTCTG 720  
Db 721 gggccatcccaattctcttacttgaatttggtctaatacaaaactagtcagggttttcca 780  
|||||  
Qy 721 GGCCATCCAAATTTCTTTACTTGAATTTGGCTAATAACAACACTAGTCAGGTTTTCGA 780  
Db 781 acctgaccgacatgaactgtcacagaaattgttccagctactatggagtgctcacaaag 840  
|||||  
Qy 781 ACCTTCACCGACATGAACGTACACAGAAATGTTTCCAGTACTTGGAGTGTCTACAAGG 840  
Db 841 atactttacaggttaagaacaaagggttgactggcctatttactatgatcaagaacatgctc 900  
|||||  
Qy 841 ATACTTTTACAGGTTAAGACAAAGGGTTGACTGGCCTATTATCTGATCAAGAACATGTC 900  
Db 901 agcaatgtctcttgtgctctaaattcttattactacaataatatttgaagatccc 960  
|||||  
Qy 901 AGCAATGCTCTTTGTGCTCTAATAATCTTATATACATAATATATTGTAAGATCC 960  
Db 961 tatagctctcttttttgagatggaggttgcctttgttgccagagtgagtgcaatgg 1020  
|||||  
Qy 961 TATAGTCTTTTTTTTGAGATGAGGTTTCCTTTTGTGCCAGGCTGGAGTGCAATGG 1020  
Db 1021 cggcatcttggctcaaccataaacctccgctccaggttcaagcaattctcctgccttagc 1080  
|||||  
Qy 1021 CGGATCTTGGCTCACATAACCTCCGCTCCAGGTTCAAGCAATTTCTCTGCTTAGC 1080  
Db 1081 ctectgagtagctgggattacaggcgtgagccactactgctgactaattttttagtttta 1140  
|||||  
Qy 1081 CTCTCTGAGTAGCTGGGATTACAGGCTGGCGCACTATGCTGACTAAATTTTGTAGTTTAA 1140  
Db 1141 gtgagacggggtttctccatgttggctcaggctggctcacaactcctgacctcagggtgat 1200  
|||||  
Qy 1141 GTAGACAGGGGGTTCTCCATGTTGGTCAGGGCTGGCTGCAAACTCTCTGACCTCAGGTGAT 1200  
Db 1201 ctcccgctcagcctcccaagtgtcgtggaattacaggcgtgagccaccacgcctggctc 1260  
|||||  
Qy 1201 CTCCCGGCTCAGCCTCCCAAGTGTGGAATTAACAGCGGTGAGCCACCGCCTGGCTG 1260  
Db 1261 gatcctatatcttagtgagacatataaacgagctctaatcacatttcacttcaagggtca 1320  
|||||  
Qy 1261 GATCCTATATCTTAGTGAACACATAAAGCGAGTCTAATTACATTTCACTTCAAGGCTCA 1320  
Db 1321 atgctattctcaataatgacaagtattttctactaaaccagaataatggtagaagatttta 1380  
|||||  
Qy 1321 ATGCTATTCTAATGACAAGTAATTTTCTACTAAACCCAGAAATTTGGTAGAGGATTTA 1380

BASE COUNT	239 a	170 c	163 g	180 t
------------	-------	-------	-------	-------

ACCESSION	U14572
NID	g551541
KEYWORDS	.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryotae; Hyperchondria; Eukaryote crown group; Metazoa/Eumycota
	group; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia;
	Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii;
	Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta;
	Primates; Catarrhini; Hominoidea; Homo.

REFERENCE	1 (bases 1 to 291)
AUTHORS	Jurka, J. and Milosavljevic, A.
TITLE	Reconstruction and analysis of human Alu genes
JOURNAL	J. Mol. Evol. 32 (2), 105-121 (1991)
MEDLINE	91178815
REFERENCE	2 (bases 1 to 291)
AUTHORS	Claverie, J.-M. and Makalowski, W.
TITLE	Alu alert
JOURNAL	Nature 371, 752-752 (1994)
MEDLINE	95021758
REFERENCE	3 (bases 1 to 291)
AUTHORS	Makalowski, W. and Claverie, J.-M.
TITLE	Direct Submission
JOURNAL	Submitted (09-SEP-1994) Wojciech Makalowski, National Center for Biotechnology Information, NIH, 8600 Rockville Pike, Bethesda, MD 20894, USA
COMMENT	CAUTION: This consensus Alu entry is provided to help identify Alu-derived sequences in newly determined nucleotide sequences.

**CAUTION:** Alu repetitive elements, unique to primates, are interspersed within the human genome with an average spacing of 4 kb. Some Alus are actively transcribed by pol III. Normal transcripts may contain Alu-derived sequences in 5' or 3' untranslated regions. However, cDNA libraries also contain partial and/or rearranged cDNAs ligated with an Alu-derived sequence in any orientation. Although Alu elements (especially situated on the complementary strand) have a great potential to create additional/alternative exons, consideration should be given to the possibility that the presence of an Alu in an open reading frame may have resulted from a cloning artifact or may be due to misinterpretation of sequencing data. This point has been overlooked on several occasions, with the consequence of erroneous Alu-derived amino acid sequences being reported.

**CAUTION:** Any significant similarity of a putative protein sequence with this entry must be taken as a warning that a portion of an Alu repeat may have been artificially included in the coding nucleotide sequence. Furthermore, the possibility exists that the full extent of an Alu sequence may not be identified by a match against these consensus sequences. The possibility of false negative results may also occur when searching a putative protein sequence against these consensus sequences.

**CAUTION:** Discovery of Alu-like sequence(s) in DNA of non-primate origin should be considered as a warning that contamination has occurred.

Various analyses (see reference 2) indicate that Alu repeats fall into 8 subfamilies. Therefore, 8 ALU WARNING consensus sequences have been constituted and added to GenBank; accession numbers U14567, U14568, U14569, U14570, U14571, U14572, U14573 and U14574.

NCBI gi: 551541  
Location/Qualifiers  
source 1..291

```

BASE COUNT      72 a   80 c   94 g   45 t
/citation=[2]
/organism="Homo sapiens"

```

Query Match 15.3%; Score 239; DB 68; Length 291;  
Best Local Similarity 92.1%; Pred. No. 2.14e-104;  
Matches 267; Conservative 0; Mismatches 22; Indels 1; Gaps 1;

Ddb	2	gcgcggcggtggctcagcctgtaatcccgagcaatttggaggcgcgagcgcgagat	61
Cp	1258	GCACGGCTGTTGGCTCAGCCCTGTAAATCCAGCACTTTGGAGGCTCAGCGGGCAGAT	11199
Ddb	62	caactgaggtcggaggttcgagaccagctcgacaacatggagaaaccccgctctcta	121
Cp	1198	CACCTGAGGTCAGGAGTTTGAGCAGCGCTGACCAACATGGAAACCCGCTCTCTACTA	11139
Ddb	122	aaatacaaaaatgagccggcggtggcgccgcatgcctgtaatccagctactcgggag	181
Cp	1138	AAACTACAAA--TTAGTCAGGCATAGTGGCGACGCCCTGTAAATCCAGCTACTCAGGAG	1080
Ddb	182	ctgaggcaggagaatcgcttgaaccgggagcgagggttgcggtgagcgagatcggc	241
Cp	1079	CTAAGGCAGCAATTCCTTGAACCTGGGAGGCGAGGTTATGTGTGAGCCACAGATCGGC	1020
Ddb	242	catcgactccagcctggcgacaagagcgaaactcgctcnaaaaaa	291
Cp	1019	CATTGCATCCAGGCTGGGCAAAAAGGCAACTTCATCTCAAAAAA	970

RESULT	4	
ID	standard; DNA; PRI; 35962 BP.	
UNBLGR1	U07561;	
AC		
23-JUL-1994	(Rel. 40, Created)	
DT	31-JAN-1996 (Rel. 46, Last updated, Version 15)	
DDT	Human ABL gene, exon 1b and intron 1b, and putative M8604 Met	
DE	protein (M8604 Met) gene, complete cds.	
DE	.	
OS	Homo sapiens (human)	
OC	Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;	
OC	Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae;	
OC	[1]	
RN	1-35962	
RP	Chisoso S.L., Bodenteich A., Wang Y., Jian L., Burian D.,	
RA	Carbtree J.S., Freeman A., Hamed O.S., McLaury H., Pan H.,	
RA	Toth S., Wang Z., Zhang G., Clifton S., Groffen J.,	
RA	Heisterkamp N., Roe B.A.;	
RT	"Sequence and analysis of the human ABL gene, the BCR gene, and	
RT	regions involved in the Philadelphia chromosomal translocation";	
RL	Genomics 27:67-82(1995).	
RN	[2]	
RN	MEDLINE; 82199444.	
RA	Groffen J., Heisterkamp N., Grosveld F., de Ven W.,	
RA	Stephenson J.R.;	
RT	"Isolation of human oncogene sequences (v-fes homolog) from a	
RT	cosmid library";	
RL	Science 216:1136-1138(1982).	
RN	[3]	
RN	MEDLINE; 91329820.	
RA	Morris C.M., Heisterkamp N., Groffen J., Fitzgerald P.H.;	
RT	"Entire ABL gene is joined with 5'-BCR in some patients with	
RT	Philadelphia-positive leukemia";	
RL	Blood 78:1078-1084(1991).	
RN	[4]	
RN	27793-28338, 28792-29265	
RP	MEDLINE; 87028219.	
RA	Shivelman E., Lifshitz B., Gale R.P., Roe B.A., Canaani E.;	
RA	"Alternative splicing of RNAs transcribed from the human abl gene	
RT	and from the bcr-abl fused gene";	
RL	Cell 47:277-284(1986).	
RN	[5]	
RN	27108-29899	
RP	MEDLINE; 88038877.	
RA		

RA Bernards A., Rubin C.M., Westbrook C.A., Paskind M., Baltimore D.;  
RT "The first intron in the human c-abl gene is at least 200  
RT kilobases long and is a target for translocations in chronic  
RT myelogenous leukemia";  
RL Mol. Cell. Biol. 7:3231-3236(1987).  
RN [6]  
RX MEDLINE; 82150208.  
RA Mount S.M.;  
RT "A catalogue of splice junction sequences";  
RL Nucleic Acids Res. 10:459-472(1982).  
RN [7]  
RX MEDLINE; 92168112.  
RA Adams M.D., Dubnick M., Kerlavage A.R., Moreno R., Kelley J.M.,  
RA Utterback T.R., Nagle J.W., Fields C., Venter J.C.;  
RT "Sequence identification of 2,375 human brain genes";  
RL Nature 355:632-634(1992).  
RN [8]  
RX MEDLINE; 90295283.  
RA Zhu Q.S., Heisterkamp N., Groffen J.;  
RT "Characterization of the human ABL promoter regions";  
RL Oncogene 5:885-891(1990).  
RN [9]  
RA Paskind M.;  
RT "Personal communication";  
RL Unpublished.  
RN [10]  
RA Tatusov R., Lipman D.J.;  
RT "Using local similarities for pattern detection in  
RT nucleotide/protein sequences";  
RL Unpublished.  
RN [11]  
RX MEDLINE; 84106827.  
RA Groffen J., Stephenson J.R., Heisterkamp N., De Klein A.,  
RA Bartram C.R., Grosfeld G.;  
RT "Philadelphia chromosomal breakpoints are clustered within a  
RT limited region, bcr, on chromosome 22";  
RL Cell 36:93-99(1984).  
RN [12]  
RP 1-35962  
RA Chissee S.L.;  
RT ;  
RL Thesis (1994), University of Oklahoma  
RN [13]  
RP 1-35962  
RA Roe B.A.;  
RT ;  
RL Submitted (08-MAR-1994) to the EMBL/GenBank/DBJ databases.  
RL Bruce A. Roe, Dept. of Chem. and Biochem., University of Oklahoma,  
RL 620 Parrington Oval, Rm 208, Norman, OK 73019, USA  
CC NCBI gi: 514262  
FH Key Location/Qualifiers  
FH  
FT source  
FT 1..35962  
FT /clone="Ca9-1"  
FT /chromosomes="9"  
FT /map="9q34"  
FT /organism="Homo sapiens"  
FT /tissue\_type="lung carcinoma"  
FT complement(448..783)  
FT repeat\_region  
FT /citation=[10]  
FT /rpt\_family="Alu"  
FT complement(478..(726.728))  
FT /citation=[10]  
FT /rpt\_family="Alu"  
FT

FT repeat\_region  
FT 1012..(1327.1338)  
FT /citation=[10]  
FT /rpt\_family="Alu"  
FT complement(1455..1786)  
FT repeat\_region  
FT /citation=[10]  
FT /rpt\_family="Alu"  
FT (1907.1915)..2184  
FT /citation=[10]  
FT /rpt\_family="Alu"  
FT 2185..2227  
FT /citation=[10]  
FT /rpt\_family="Alu"  
FT complement(2619..2894)  
FT repeat\_region  
FT /citation=[10]  
FT /rpt\_family="Alu"  
FT 3454..3774  
FT /citation=[10]  
FT /rpt\_family="Alu"  
FT (4049.4097)..(4411.4433)  
FT repeat\_region  
FT /citation=[10]  
FT /rpt\_family="Alu"  
FT complement((5798.5804)..6146)  
FT repeat\_region  
FT /citation=[10]  
FT /rpt\_family="Alu"  
FT complement(6412..(6677.6691))  
FT repeat\_region  
FT /citation=[10]  
FT /rpt\_family="Alu"  
FT 6510..6820  
FT /citation=[10]  
FT /rpt\_family="Alu"  
FT complement((6980.6990)..(7289.7340))  
FT repeat\_region  
FT /citation=[10]  
FT /rpt\_family="Alu"  
FT <8604..8719  
FT /number=1  
FT /product="unknown"  
FT join(8604..8719,10298..10399,12388..12433,14134..14199,  
FT 15672..15740,16939..17115,17862..17990,18503..18583)  
FT /evidence=EXPERIMENTAL  
FT /note="putative 8604 MET gene, 5' to the ABL gene; NCBI gi  
FT :  
FT 514266"  
FT /citation=[6]  
FT /citation=[7]  
FT /codon\_start=1  
FT /product="unknown"  
FT /db\_xref="PID:g514266"  
FT /translation="MEMRLPVARKELSERLCGRDTKKHLVPGDTITTTGEMRGHCTY  
FT M  
FT GEEKLIASVAGSVERNKLCVKALKTRYIGEVGDIVVGRI TERRRSAEDELAMRGFL  
FT Q  
FT EGDLSAEVQAVFSDGAVSLHTRSLIKYKLGQGVLVQVSPSLVKRQKTHFDLPCCAS  
FT V  
FT ILGNNGFTWITYPTPEKKEEAGGFTANLEPVSADREVISRINCCIISLVYQRMMLYD  
FT T  
FT SILCYEASIPHQIKDILAKPEIMEEIVNMETRQRILLEQEG"  
FT complement(9524..9833)  
FT /citation=[10]  
FT /rpt\_family="Alu"  
FT 10298..10399  
FT /number=2  
FT /product="unknown"  
FT complement(10663..10902)  
FT repeat\_region  
FT /citation=[10]



```

REFERENCE 10 (sites)
AUTHORS Tatusov,R. and Lipman,D.J.
TITLE Using local similarities for pattern detection in
nucleotide/protein sequences
JOURNAL Unpublished
REFERENCE 11 (sites)
AUTHORS Groffen,J., Stephenson,J.R., Heisterkamp,N., de Klein,A.,
Bartram,C.R. and Grosveld,G.
TITLE Philadelphia chromosomal breakpoints are clustered within a limited
region, bcr, on chromosome 22
JOURNAL Cell 36 (1), 93-99 (1984)
MEDLINE 84106827
REFERENCE 12 (bases 1 to 35962)
AUTHORS Chisoe,S.L.
TITLE Sequence of the human abl and bcr genes
JOURNAL Thesis (1994) University of Oklahoma
REFERENCE 13 (bases 1 to 35962)
AUTHORS Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (08-MAR-1994) Bruce A. Roe, Dept. of Chem. and Biochem.,
University of Oklahoma, 620 Parrington Oval, Rm 208, Norman, OK
73019, USA
COMMENT NCBI gi: 514262
FEATURES
source
Location/Qualifiers
1..35962
/clone="Ca9-1"
/chromosomes="9"
/map="9q34"
/organism="Homo sapiens"
/tissue_type="lung carcinoma"
/citation=[2]
repeat_region
/citation=[3]
complement(448..783)
/rpt_family="Alu"
repeat_region
/citation=[10]
complement(478..726.728))
/rpt_family="Alu"
repeat_region
1012..(1327.1338)
/citation=[10]
/rpt_family="Alu"
repeat_region
/citation=[10]
complement(1455..1786)
/rpt_family="Alu"
repeat_region
/citation=[10]
complement(1907.1915)..2184
/rpt_family="Alu"
repeat_region
2185..2227
/rpt_family="Alu"
repeat_region
/citation=[10]
complement(2619..2894)
/rpt_family="Alu"
repeat_region
3454..3774
/rpt_family="Alu"
repeat_region
/citation=[10]
complement(4049.4097)..(4411.4433)
/rpt_family="Alu"
repeat_region
/citation=[10]
complement(5798.5804)..6146)
/rpt_family="Alu"
repeat_region
/citation=[10]
complement(6412..(6677.6691))

```

```

repeat_region
/rpt_family="Alu"
/citation=[10]
6510..6820
/rpt_family="Alu"
repeat_region
/citation=[10]
complement((6980.6990)..(7289.7340))
/rpt_family="Alu"
CDS
/citation=[10]
join(8604..8719,10298..10399,12388..12433,14134..14199,
15672..15740,16939..17115,17862..17990,18503..18583)
/note="putative 8604 MET gene, 5' to the ABL gene; NCBI
gi: 514266"
/codon_start=1
/evidence=experimental
/citation=[6]
/citation=[7]
/product="unknown"
/db_xref="PID:G514266"
/translation="MEMRLPVARKPLSERLGRDTKKHLVPGDTITDTGFMRGHCTY
MGEKLIASVAGSVERNKLIQKALKTRYIGVGDIIVGRITERRRAEELAMRGF
LQEGDLISAEQVAFSDCAVSLHTSLKYKLGQGLVQVPSLVKROKTHFDLPFG
ASVILNNGCFIWIYPTPEHKEEAGGFANLEPVSILADREVISRINCIISLVTRQWM
LYDTSILYCEASLPHQIKDILKPEIMEEIVMETRQLLEQEG"
exon
<8604..8719
/number=1
/product="unknown"
repeat_region
complement(9524..9833)
/rpt_family="Alu"
exon
/citation=[10]
10298..10399
/number=2
/product="unknown"
repeat_region
complement(10663..10902)
/rpt_family="Alu"
repeat_region
/citation=[10]
complement(10906..10991)
/rpt_family="Alu"
exon
/citation=[10]
12388..12433
/number=3
/product="unknown"
repeat_region
complement(12599..12899)
/rpt_family="Alu"
repeat_region
/citation=[10]
13035..(13382.13389)
/rpt_family="Alu"
repeat_region
/citation=[10]
complement((13422.13425)..13719)
/rpt_family="Alu"
repeat_region
/citation=[10]
complement(13922..14172)
/rpt_family="Alu"
exon
/citation=[10]
14134..14199
/number=4
/product="unknown"
repeat_region
(14674.14676)..(15029.15040)
/rpt_family="Alu"
repeat_region
/citation=[10]
complement((14984.15063)..(15359.15379))
/rpt_family="Alu"
repeat_region
/citation=[10]
complement((15133.15140)..15489)
/rpt_family="Alu"

```



/citation=[10]

repeat\_region

15367..15603

/rpt\_family="Alu"

... Note: remainder of annotations omitted.

Query Match 15.3%; Score 238; DB 49; Length 35962;  
Best Local Similarity 91.3%; Pred. No. 7.33e-104;  
Matches 263; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Db 15175 ttttttgacgagagtttctactttgttcccgaggtgagtggaatgttgatctcg 15234

Qy 971 TTTTITGAGATGAGTTTCCTTTTGTGCGCAGGCTGGAGTGCATATGGCGCATCTTG 1030

Db 15235 gctcactgaacctctgctccgggttcaagcattctcctgctcagctcctcagta 15294

Qy 1031 GCTCACCATAACCTCGGCTCCAGGTTCAAGCAATTCCTCGCTTAGCCTCCTGAGTA 1090

Db 15295 gctgggattcacagcagtcgcccactatgccgactaattttgtacttttagtagagcgg 15354

Qy 1091 GCTGGCAATACAGGCTCGCCACTATGCTGACTAATTTGTAGTTTATAGACACGG 1150

Db 15355 ggtttctcatgttggtcaggtggtctagaactcctgacctcaggtgacgtcct 15414

Qy 1151 GGTTCCTCATGTTGTCAGCTGGTCTCAAACTCTGACCTCAGGTGATCGCCGCT 1210

Db 15415 cagctcccaagtgtgggtgtacaggtgtgagccaccgacctgpc 15462

Qy 1211 CAGCCTCCAAAGTGTGAATTACAGGCGTGAGCCACCGCCTGCG 1258

RESULT 6  
LOCUS HSABLGR1 35962 bp DNA PRI 31-JAN-1996  
DEFINITION Human ABL gene, exon 1b and intron 1b, and putative M8604 Met protein (M8604 Met) gene, complete cds.  
ACCESSION U07561  
NID g514262  
KEYWORDS  
SEGMENT  
SOURCE 1 of 3 human.  
ORGANISM Homo sapiens  
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 35962)  
AUTHORS Chisoe, S.L., Bodenteich, A., Wang, Y., Jian, L., Burian, D., Carbtree, J.S., Freeman, A., Hamed, O.S., McLauri, R., Pan, H., Toth, S., Wang, Z., Zhang, G., Clifton, S., Groffen, J., Heisterkamp, N. and Roe, B.A.

TITLE Sequence and analysis of the human ABL gene, the BCR gene, and regions involved in the Philadelphia chromosomal translocation  
JOURNAL Genomics 27 (1), 67-82 (1995)  
MEDLINE 95394474  
REFERENCE 2 (sites)  
AUTHORS Groffen, J., Heisterkamp, N., Grosveld, F., Van de Ven, W. and Stephenson, J.R.

TITLE Isolation of human oncogene sequences (v-fes homolog) from a cosmid library

JOURNAL Science 216 (4550), 1136-1138 (1982)  
MEDLINE 82199444  
REFERENCE 3 (sites)  
AUTHORS Morrie, C.M., Heisterkamp, N., Groffen, J. and Fitzgerald, P.H.  
TITLE Entire ABL gene is joined with 5'-BCR in some patients with Philadelphia-positive leukemia  
JOURNAL Blood 78 (4), 1078-1084 (1991)

MEDLINE 91329820  
REFERENCE 4 (bases 27793 to 28338; 28792 to 29265)  
AUTHORS Shivelman, E., Lifshitz, B., Gale, R.P., Roe, B.A. and Canaani, E.  
TITLE Alternative splicing of RNAs transcribed from the human abl gene and from the bcr-abl fused gene  
JOURNAL Cell 47 (2), 277-284 (1986)  
MEDLINE 87028219  
REFERENCE 5 (bases 27108 to 29899)  
AUTHORS Bernards, A., Rubin, C.M., Westbrook, C.A., Paskind, M. and Baltimore, D.

TITLE The first intron in the human c-abl gene is at least 200 kilobases long and is a target for translocations in chronic myelogenous leukemia

JOURNAL Mol. Cell. Biol. 7 (9), 3231-3236 (1987)  
MEDLINE 88038877  
REFERENCE 6 (sites)  
AUTHORS Mount, S.M.  
TITLE A catalogue of splice junction sequences  
JOURNAL Nucleic Acids Res. 10 (2), 459-472 (1982)  
MEDLINE 82150208  
REFERENCE 7 (sites)  
AUTHORS Adams, M.D., Dubnick, M., Kerlavage, A.R., Moreno, R., Kelley, J.M., Utterback, T.R., Nagle, J.W., Fields, C. and Venter, J.C.

TITLE Sequence identification of 2,375 human brain genes  
JOURNAL Nature 355 (6361), 632-634 (1992)  
MEDLINE 92168112  
REFERENCE 8 (sites)  
AUTHORS Zhu, Q.S., Heisterkamp, N. and Groffen, J.

TITLE Characterization of the human ABL promoter regions  
JOURNAL Oncogene 5 (6), 885-891 (1990)  
MEDLINE 90295283  
REFERENCE 9 (sites)  
AUTHORS Paskind, M.

TITLE Personal communication  
JOURNAL Unpublished  
REFERENCE 10 (sites)  
AUTHORS Tatusov, R. and Lipman, D.J.

TITLE Using local similarities for pattern detection in nucleotide/protein sequences  
JOURNAL Unpublished  
REFERENCE 11 (sites)  
AUTHORS Groffen, J., Stephenson, J.R., Heisterkamp, N., de Klein, A., Bartram, C.R. and Grosveld, G.

TITLE Philadelphia chromosomal breakpoints are clustered within a limited region, bcr, on chromosome 22  
JOURNAL Cell 36 (1), 93-99 (1984)  
MEDLINE 84106827  
REFERENCE 12 (bases 1 to 35962)  
AUTHORS Chisoe, S.L.

TITLE Sequence of the human abl and bcr genes  
JOURNAL Thesis (1994) University of Oklahoma  
REFERENCE 13 (bases 1 to 35962)  
AUTHORS Roe, B.A.

TITLE Direct Submission  
JOURNAL Submitted (08-MAR-1994) Bruce A. Roe, Dept. of Chem. and Biochem., University of Oklahoma, 620 Parrington Oval, Rm 208, Norman, OK 73019, USA

COMMENT NCBI gi: 514262  
FEATURES Location/Qualifiers  
source 1..35962  
/clone="Ca9-1"  
/chromosome="9"  
/map="9q34"  
/organism="Homo sapiens"



	/tissue_type="lung carcinoma" complement(448..783) /citation=[10] /rpt_family="Alu" complement(478..(726.728)) /citation=[10] /rpt_family="Alu" 1012..(1327.1338) /citation=[10] /rpt_family="Alu" complement(1455..1786) /citation=[10] /rpt_family="Alu" (1907.1915)..2184 /citation=[10] /rpt_family="Alu" 2185..2227 /citation=[10] /rpt_family="Alu" complement(2619..2894) /citation=[10] /rpt_family="Alu" 3454..3774 /citation=[10] /rpt_family="Alu" (4049.4097).. /citation=[10] /rpt_family="Alu" complement(5798.5804)..6146) /citation=[10] /rpt_family="Alu" complement(6412..(6677.6691)) /citation=[10] /rpt_family="Alu" 6510..6820 /citation=[10] /rpt_family="Alu" complement((6980.6990).. /citation=[10] /rpt_family="Alu" <8604..8719 /number=1 /product="unknown"
CDS	join(8604..8719,10298..10399,12388..12433,14134..14199, 15672..15740,16939..17115,17862..17990,18503..18583) /evidence=experimental /note="putative 8604 MET gene, 5' to the ABL gene; NCBI gi: 514266" /citation={6} /citation={7} /codon_start=1 /product="unknown" /db_xref="PID:g514266" /translations="MEMPLPVARKP LSR LGRDTKKHLVVP EDTITTTDGFMRGHGYT MGEKLLTSVAGSVEVRNKLVCKALKTTRYIGEVGDIVGRI TERRQSAEDFLAMRCF LQEGDLISAEQAVFS DGNLSLTRSYKLGCGGVQSVLRQRKTHFDLPCCG ASVI LGNNGTWITYTPPEHEEAGGF TANLEPSVLADREVISRLNCTIISLVTRQM LYDYSILCYEAESLP HQIKLKPEIMEEI VME TRQLLEQG"
repeat_region	complement(9524..9833) /citation=[10] /rpt_family="Alu" 10298..10399 /number=2 /product="unknown"
exon	

[illegible]

```

QY 1211 CAGCCTCCCAAGTGTGGAATTACAGCGGTGAGCCACCGCCCTGGC 1258
|||||
RESULT 7
LOCUS HSATPCP1 9457 bp DNA PRI 19-JUL-1993
DEFINITION H.sapiens gene for mitochondrial ATP synthase c subunit (P1 form).
ACCESSION X69907
NID 938429
KEYWORDS ATP synthase; ATP synthase c subunit.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa/Eumycota group;
Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata;
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 9457)
AUTHORS Walker, J.E.
TITLE Direct Submission
JOURNAL Submitted (29-DEC-1992) to the EMBL/GenBank/DBJ databases. J.E.
Walker, M.R.C. Lab. of Molecular Biology, Hills Road, Cambridge CB2
2QH., UK
REFERENCE 2 (bases 1 to 9457)
AUTHORS Dyer, M.R. and Walker, J.E.
TITLE Sequences of members of the human gene family for the c subunit of
mitochondrial ATP synthase
JOURNAL Biochem. J. 293 (Pt 1), 51-64 (1993)
MEDLINE 93319529
COMMENT NCBI gi: 38429
FEATURES
source
1..9457
/organism="Homo sapiens"
/cell_type="T-cell"
/cell_line="AT5B1"
/clone_lib="lambda 2001"
1..224
/rpt_family="Alu"
273..442
/rpt_family="Alu"
606..904
/rpt_family="Alu"
1449..1768
/rpt_family="Alu"
1769..2111
/rpt_family="Alu"
2627..2945
/rpt_family="Alu"
2646..3233
/rpt_family="Alu"
4041..4046
/rpt_family="Alu"
4100..4792
/rpt_family="Alu"
4538..4542
CAAT_signal
4641..4645
CAAT_signal
4671..4675
CAAT_signal
4697..4701
CAAT_signal
4708..4713
TATA_signal
mRNA
join(4733..4842,5341..5388,6304..6381,7088..7266,
7587..7802)
/gene="P1 gene for c subunit of human mitochondrial ATP
synthase"
prim_transcript 4733..7802
/gene="P1 gene for c subunit of human mitochondrial ATP

```

```

synthase"
4733..4842
/gene="P1 gene for c subunit of human mitochondrial ATP
synthase"
/number=1
4843..5340
/number=1
5002..5007
TATA_signal
5279..5280
exon
5341..5388
/gene="P1 gene for c subunit of human mitochondrial ATP
synthase"
/number=2
join(5350..5388,6304..6381,7088..7266,7587..7701)
/gene="P1 gene for c subunit of human mitochondrial ATP
synthase"
/note="NCBI gi: 38430"
/codon_start=1
/db_xref="PID:Q38430"
/translation="MQTAGALFISPALRCRGLIRPVSAFLNSPVSNSKQPSYSN
FPLQVAREFQTSVSRDIDTAKEIGAGAAATVGAGGAGICTVFGSLIIGYARNPS
LQQLFVAILGLFALSEAMGLFCLMVAFLILFAM"
5389..6303
/number=2
6304..6381
/gene="P1 gene for c subunit of human mitochondrial ATP
synthase"
/number=3
6382..7087
/number=3
6418..6699
/rpt_family="Alu"
7088..7266
/gene="P1 gene for c subunit of human mitochondrial ATP
synthase"
/number=4
join(7154..7266,7587..7698)
/gene="P1 gene for c subunit of human mitochondrial ATP
synthase"
7267..7586
/number=4
7587..7802
/gene="P1 gene for c subunit of human mitochondrial ATP
synthase"
/number=5
7780..7785
/gene="P1 gene for c subunit of human mitochondrial ATP
synthase"
8879..9176
/rpt_family="Alu"
BASE COUNT 2505 a 2208 c 2318 g 2426 t
ORIGIN
Query Match 15.2%; Score 237; DB 49; Length 9457;
Best Local Similarity 90.7%; Pred. No. 2.5le-103;
Matches 264; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
Db 8878 agccagcatggtggctcacgcctgtaatcccagcacttggaggagtgagccgcgga 8937
|||||
Cp 1259 AGCCAGCGTGGTGGCTCAGCGCTGTATTCCAGCAGCTTTGGGAGGCTGAGCGGCGAGA 1200
|||||
Db 8938 tcacctgaggtcaggagttcgagaccagcctgacaacatggagaaacccagctctact 8997
|||||
Cp 1199 TCACCTGAGGTGAGGAGTTTCAGACACCGCTGACCAACATGAGAACCCCGCTCTACT 1140

```

```

NCBI gi: 34286
FEATURES
    source             location/Qualifiers
    1..6901
    /organism="Homo sapiens"
    /clone_lib="(lambda)charon30"
    709..716
    CAAT_signal
    TATA_signal
    786..790
    prim_transcript
    809..5052
    mRNA
    4458..5052
    809..990
    exon
    /number=1
    1001..1002
    CDS
    join(837..990,1724..1880,1960..2075,2170..2265,2349..2573,
    4458..5032)
    /note="NCBI gi: 34287"
    /codon_start=1
    /product="lecithin-cholesterol acyltransferase (LCAT)"
    /db_xref="pid:g34287"
    /translation="MGPPGSPQWVLTLLGLLPPPAAPFWLLNLVLPFHTTKAELSN
    TRPVVLVPGCLGQLQLEAKDLPDVVNMWYKRTEDFTIWLNNFLPLGWDCHIND
    TRVNVNSSLGSLVSNAPGVQIRVPEFGKTYSEYLDSSKLAGYLHTLVQNLVNWNGVVR
    ETRVAPVDMRLSEPCCQEEYKRLAGLVEEMHAAVKPVLIGHSLGCLHLVFLLRG
    EQWAKDFDGEISLGAPEGGCSIKPMVLASGNDGQIPIMSSIKGLKEEQRIITTPSPW
    FFSMAWPEHDHFVISTESFNVTGRQRFADLHFEECGYMMIQSRDILLAGLAPGVVR
    VYCLYGVLPTPTIYDHGEPYTDPPGVLYEDGDDVATRS TELCGLWQGRQPQVH
    LIPLHGILQHNLNVNFTSLTLEHINAILLGATROGPPASPTASPEPPPE"
    991..1723
    intron

```

RESULT	9	HSTMP04	6905 bp	DNA	PRI	02-JUL-1995
LOCUS						



```
/note="putative Spl"
820..1131
/note="Alu repeat"
1134..1296
/note="Alu repeat"
1350..1482
/note="Alu repeat"
1576..1581
/note="putative Spl"
2251..3970
/note="CpG island"
2267..2272
RBS 2308..2313
/note="putative Spl"
<2310..2400
/number=1
join(2345..2400,2609..2696,2782..2879,2991..3131,
3395..3510,3615..3673,4145..4296,4422..4533)
/note="NCBI gi: 406227"
/codon_start=1
/product="protease-like subunit MECL-1"
/db_xref="PID:g406227"
/translation="MLRPALEPRGGSFENCQRNASLERVLPGLKVPHARKTGTIAG
LVFGVILGADTRATNDVVDKCEKIHFIAPKIYCGGVAADAEITTRVASKM
ELHALSTGRFPAVTRILRLQTLFRYQGHVGASLIVGVDLTGPGLYGVHPGYSR
LPFTALGSDAALAVLEDLRFQPNMTLEAAQGLLVEAVTAGILGDLGGGNVACVIT
KTKAKLLRTLSSTPEPKRSGRYHEVPQTAVLTQVXP LTLVETVQAVEVE"
2345..2400
/misc_feature
/label=test
intron 2401..2608
/number=1
RBS 2419..2424
/note="putative Spl"
exon 2609..2696
/number=2
intron 2697..2781
/number=2
exon 2782..2879
/number=3
intron 2880..2990
/number=3
exon 2991..3131
/number=4
intron 3132..3394
/number=4
RBS 3208..3212
RBS 3376..3381
exon 3395..3510
/number=5
RBS 3428..3433
intron 3511..3614
/note="putative Spl"
exon 3615..3673
/number=5
intron 3674..4144
/number=6
exon 4145..4296
/number=7
intron 4297..4421
/number=7
polyA_signal 4363..4568
```

```
exon 4422..>4568
/number=8
misc_feature 4422..4533
/label=test2
repeat_unit 5560..6000
/note="Alu repeat"
RBS 5942..5948
/note="putative Spl"
repeat_unit 6037..6326
/note="Alu repeat"
TATA_signal 7202..7207
prim_transcript 7232..9523
/note="C"
prim_transcript 7232..9484
/note="B"
prim_transcript 7232..9210
/note="A3"
prim_transcript 7232..9220
/note="A2"
prim_transcript 7232..9216
/note="A1"
exon 7232..7293
/number=1
misc_feature 7242..7293
/label=et
join(7242..7293,7893..7996,8095..8174,8287..8368,
8506..8686,8774..8907,8998..9159)
/note="NCBI gi: 406228"
/codon_start=1
/product="chymotrypsin-like protease CTRL-1"
/db_xref="PID:g406228"
/translation="MLLSLTSLVLGSSWGGCIPAIKPAISFSQRIWGENAVLGS
WPQVSIQDSSGFEFCGSLISQSWVTAAHCNVSPGRHFVILGEYDRSSNAEP LQVL
SVSRATHPSWNSTTMNDVTILKLSPAOYTTIRISPVCLASNEALTEGILCVTTGW
GRLSGCVNPVPAHLQQLVALPLVTWQCQYWGSSITDSMICAGGASSCGGSGGL
VQKGNVTWILGIVSWGTRCNVRAPVYTRVSKFTWQINQVIAYN"
sig_peptide join(7242..7293,7893..7894)
exon 7242..7293
/note="signal peptide"
/number=1
intron 7294..7892
/note="signal peptide"
/number=1
exon 7893..7996
/number=2
exon 7893..7894
/note="signal peptide"
/number=2
misc_feature 7895..7939
/note="activation peptide"
misc_feature 7940..7996
/label=et1
mat_peptide join(7940..7996,8095..8174,8287..8368,8506..8686,
8774..8907,8998..9159)
/product="chymotrypsin-like protease CTRL-1"
intron 7997..8094
/number=2
exon 8095..8174
/number=3
intron 8175..8286
/number=3
exon 8287..8368
/number=4
intron 8369..8505
```

```

/number=4
8506..8686
/number=5
8687..8773
/number=5
8774..8907
/number=6
8908..8997
/number=6
8998..9159
/label=el
8998..9159
/label=e
8998..9216
/note="number 7A1"
8998..9523
/note="number 7C"
8998..9484
/note="number 7B"
8998..9210
/note="number 7A3"
8998..9220
/note="number 7A2"
9190..9195
/note="number A"
complement(9415..11763)
/number=2
prim_transcript
complement(9415..13863)
misc_feature
9415..9523
/note="overlapping transcription units"
polyA_signal
9462..9467
/note="number B"
polyA_signal
9507..9512
...
Note: remainder of annotations omitted.

```

Query Match	14.9%;	Score 232;	DB 52;	Length 13863;
Best Local Similarity	88.9%;	Pred. No. 1.19e-100;		
Matches	265;	Conservative	0;	Mismatches 33;
			Indels	0;
			Gaps	0;

Db	513	agccttggtgcggtggctcatgcctgtaatcccagcaacctggagagcccgaggtggcgcca	572
Cp	1259	AGCGAGCGCTGCTGGCTCACGCGCTGTAATTCACGACACTTTGGGAGGCTCAGGCGGCAGA	1200
Db	573	tcacttgaggtcagaagtttgagaccacctgaccaaatggagaaccccgctctctaet	632
Cp	1199	TCACCTGAGGTCAGGAGTTTGAGACGAGCTGACCAATCGAGAAACCCGCTCTACT	1140
Db	633	aaaaatcacaaattagccgggcatggtggcgcatgcctgtaatcccagctactcgggagg	692
Cp	1139	AAACTACAAAATAGTCAGGCATAGTGGCGGACGCGCTGTAATCCACGCTACTCAGGAGG	1080
Db	693	ctgagcgaggagaattgttgaacctgggaggtggagctggcgtgagccaagatcacac	752
Cp	1079	CTAAGCGAGGAGAAATGCTTGAACTCGGAGCGGAGGTTATGGTGAGCGCAAGTCGCC	1020
Db	753	tattgcactccagcctgggcaacaagagcaaaactccgctctcaaaaaaaaatttat	810
Cp	1019	CATTGCRCTCAGCCTGGGCAACAAAAGGAACTCCATCTCTCAAAAAAAGACCTAT	962

RESULT	11
ID	HS169F7B
AC	269650;
	standard; DNA; PRI; 16851 BP.

DT	22-FEB-1996 (Rel. 46, Created)
DT	22-FEB-1996 (Rel. 46, Last updated, Version 1)
DE	Human DNA sequence from cosmid L69F7B, Huntington's Disease Region,
DE	chromosome 4p16.3 contains Huntington Disease (HD) gene.
DE	4p16.3; Huntington Disease (HD) gene.
KW	Homo sapiens (human)
OC	Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC	Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homidae.
RN	[1]
RP	1-16851
RT	Mungall A, Odell C.;
RA	;
RL	Submitted (16-FEB-1996) to the EMBL/GenBank/DBJ databases.
RL	Sanger Centre, Hinxton, Cambridgeshire, CB10 1RQ, UK. E-mail
RL	enquires: humquery@sanger.ac.uk
CC	IMPORTANT:
CC	This sequence is not the entire insert of clone L69F7B. It may be
CC	shorter because we only sequence overlapping sections once, or
CC	longer because we arrange for a small overlap between neighbouring
CC	submissions.
CC	This sequence has been finished according to sequence map criteria
CC	as follows. An attempt is made to resolve all sequencing problems,
CC	such as compressions and repeats, but not necessarily within known
CC	annotated human repeat sequence elements (e.g. Alu). Where the
CC	sequence is ambiguous, there is an annotation using the "unsure"
CC	feature key.
CC	The sequence from clone L161A8 has been finished in more than one
CC	contig.
CC	This sequence (L161A8B) is separated from the preceeding one
CC	(L161A8A) by
CC	a gap of 1100bp.
CC	The true left end of clone L161A8 is at 16748.
CC	L69F7 is from cosmid library LA04NC01 constructed at the
CC	Human Genome Center, Los Alamos National Laboratory, NM 87545
CC	under the auspices of the U.S. Department of Energy.
CC	The library was constructed using flow-sorted human chromosome 4
CC	from a Hamster-Human hybrid cell line (UV20HL21-27) containing
CC	human chromosomes 4, 8 and 21.
CC	VECTOR: sCos1
CC	L69F7 is contained in a clone contig spanning ~2Mb which is
CC	described in Baxendale et al, Nature Genetics 4 ( 1993 ) 181-
CC	186. See also Myers et al, Cytogenet Cell Genet. 66 ( 1994 )
CC	218-230.
FH	Key
FH	Location/Qualifiers
FT	source
FT	1..16851
FT	/organism="Homo sapiens"
FT	/chromosome="4"
FT	/map="4p16.3"
FT	/clone="L69F7B"
FT	/clone.lib="LA04NC01"
FT	/cell_line="UV20HL21-27"
FT	complement (Join (<1187..1313, 7404..7606, 8663..8815, 10383,
FT	10459, 12716..12836, 14946..>15092))
FT	/product="Huntington Disease (HD) gene"
FT	/partial
FT	/codon_start=2
FT	/db_xref="PID:e"
FT	1316..1353
FT	/note="19 copies of 2 mer 84 % conserved"
FT	4203..4240
FT	/note="19 copies of 2 mer 84 % conserved"
FT	2282..2392
FT	/note="WSTC element fragment"

FT repeat\_region 2313..2366  
/note="MLT1A element fragment"  
FT repeat\_region 2436..2572  
/note="MLT1A element fragment"  
FT repeat\_region 2497..2571  
/note="MLT1C element fragment"  
FT repeat\_region 3026..3064  
/note="L1 element fragment"  
FT repeat\_region 3109..3155  
/note="L1 element fragment"  
FT repeat\_region 3831..4026  
/note="MER1B element fragment"  
FT repeat\_region 4022..4128  
/note="MSTA element fragment"  
FT repeat\_region 4241..4452  
/note="MSTA element fragment"  
FT repeat\_region 4450..4532  
/note="MER1B element fragment"  
FT repeat\_region 5242..5435  
/note="MER3 element fragment"  
FT repeat\_region 81..374  
/note="Alu repeat: matches 1..308 of consensus"  
FT repeat\_region 2666..2889  
/note="Alu repeat: matches 1..236 of consensus"  
FT repeat\_region 3218..3504  
/note="Alu repeat: matches 1..308 of consensus"  
FT repeat\_region 3505..3674  
/note="Alu repeat: matches 122..308 of consensus"  
FT repeat\_region 6004..6306  
/note="Alu repeat: matches 1..308 of consensus"  
FT repeat\_region 10993..11279  
/note="Alu repeat: matches 1..308 of consensus"  
FT repeat\_region 11703..11994  
/note="Alu repeat: matches 1..308 of consensus"  
FT repeat\_region 16105..16396  
/note="Alu repeat: matches 1..308 of consensus"  
FT repeat\_region 16691..16851  
/note="Alu repeat: matches 1..177 of consensus"  
FT repeat\_region 15550..15627  
/note="Alu repeat: matches 82..1 of consensus"  
FT repeat\_region 15332..15549  
/note="Alu repeat: matches 308..75 of consensus"  
FT repeat\_region 13831..14121  
/note="Alu repeat: matches 308..1 of consensus"  
FT repeat\_region 12159..12446  
/note="Alu repeat: matches 308..1 of consensus"  
FT repeat\_region 7098..7241  
/note="Alu repeat: matches 158..1 of consensus"  
FT repeat\_region 6804..7095  
/note="Alu repeat: matches 308..1 of consensus"  
FT repeat\_region 6645..6803  
/note="Alu repeat: matches 308..149 of consensus"  
FT repeat\_region 4563..4856  
/note="Alu repeat: matches 308..1 of consensus"

FT repeat\_region 1682..1966  
/partial  
/note="Alu repeat: matches 308..1 of consensus"  
FT repeat\_region 16851 BP; 5048 A; 3835 C; 3519 G; 4449 T; 0 other;  
Matches 261; Conservative 0; Mismatches 29; Indels 0; Gaps 0;  
Query Match 14.9%; Score 232; DB 10; Length 16851;  
Best Local Similarity 90.0%; Pred. No. 1.19e-100;  
Matches 261; Conservative 0; Mismatches 29; Indels 0; Gaps 0;  
Db 11704 gccaggcatggtgtcacacctgtatccagcactttggggggtcgaggaggcgagat 11763  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||  
Cp 1258 GCCAGCGGTGGCTCAGCGCTGTATTCCAGCACTTTGGAGCGCTGAGCGGCGAGAT 1199  
Db 11764 cacttgaggttggagtttgagaccagctgacaaacatggagaaacccctgtctacta 11823  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||  
Cp 1198 CACCTGAGGTGAGGAGTTTGAGACCAGCTGACCAACATGGAGAAACCCGCTCTACTA 1139  
Db 11824 aaaaatacaaaattagccgggcatggtgtgcctgtaatccagctactcaggagc 11883  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||  
Cp 1138 AAACACAAAATTAGTCAGGCATAGTCGCCAGCGCTGTAAATCCAGCTACTCAGAGGC 1079  
Db 11894 tggaggcaggagaattgcttgacccggagggtggaggttgggtgagctgagatcatgcc 11943  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||  
Cp 1078 TAGGCGGAGCAATTCCTTGACCTGGGAGCGGAGGTTATGCTGAGCCAGATCGCGGC 1019  
Db 11944 attgcactccagctgggcaacaagagcaaaactctgtctcaaaaaaaa 11993  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||  
Cp 1018 ATTGCACCTCCAGCGCTGGGCAACAAAGAGCAAACTCCATCTCAAAAAAAA 969

## RESULT 12

ID HSF1AG6PD standard; DNA; PRI; 219447 BP.  
AC 144140;  
DT 29-FEB-1996 (Rel. 47, Created)  
DT 29-FEB-1996 (Rel. 47, Last updated, Version 2)  
DE Homo sapiens chromosome X region from filamin (FLN) gene to  
DE glucose-6-phosphate dehydrogenase (G6PD) gene, complete cds's.  
KW 1A gene; 2.19 gene; ABP-280 gene; actin-binding protein;  
KW DNase I-like protein gene; DNLI1L gene; EMD gene; emerin;  
KW emery-dreifuss syndrome; filamin; FLN gene; G4.5 gene; G4.8 gene;  
KW G6PD gene; gdi gene; GdX gene; glucose-6-phosphate dehydrogenase;  
KW P3 gene; QW gene; STA gene; XAP-1 gene; XAP-2 gene; XAP-4 gene;  
KW XAP-5 gene; XAP-7 gene.  
OS Homo sapiens (human)  
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;  
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.  
RN [1]  
RP 1-219447  
RA Chen E.Y., Zollo M., Mazzarella R.A., Ciccodicola A., Chen C.N.,  
RA Zuo L., Heiner C., Burrough F.W., Ripetto M., Schlessinger D.,  
RA D'Urso M.;  
RT "Thirteen known and six candidate genes in 219.4kb of high GC DNA  
RT between the human RCP/GCP and G6PD loci";  
RL Unpublished.  
CC Submitted by: Ellison Chen, Advanced Center for Genetic Technology,  
CC Applied Biosystems Division of Perlin Elmer Corp., 850 Lincoln  
CC Center Drive, Foster City, CA 94044 USA and David Schlessinger,  
CC Department of Molecular Microbiology and Center for Genetics in  
CC Medicine Washington University School of Medicine, St. Louis MO  
CC 63110 USA e-mail: ellison@genseq.apltdbio.com and  
CC david@genetics.wustl.edu Note: Gene predictions were accomplished  
CC with runs of Grail versions 1.1 and 1.2, coupled with fasta and  
CC blastx comparisons to genbank & non-redundant peptide libraries.

Aug 27 08:17

US-08-231-565A-1.rge

31

CC Repeat analysis was accomplished via censor. NCBI gi: 1203968

Key	Location/Qualifiers
source	1..219447 /organism="Homo sapiens" /sequenced_mol="DNA" 4..223 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sp or Alu-Sq" 226..503 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J or Alu-S" 641..931 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J or Alu-S" 1023..1313 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sx" complement(1446..1536) /evidence=EXPERIMENTAL /map="X"
repeat_region	/rpt_family="Alu-J" 3095..3426 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="CpG Island" complement(3663..3952) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sb0 or Alu-Sb1" 4912..5046 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J or Alu-S" 5047..5333 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sb0 or Alu-Sb1" 5346..5628 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J or Alu-S" 5632..5805 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sq" 6363..17698 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="11Kb repeat1" 6369..6633 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sx" complement(7121..7409) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sp" complement(7358..7432) /evidence=EXPERIMENTAL

Aug 27 08:17

US-08-231-565A-1.rge

32

repeat_unit	/map="X" /rpt_family="Alu-Sb0 or Alu-Sb1" complement(7426..7701) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J" 8306..8592 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sx" complement(8863..9029) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J or Alu-S" complement(9031..9317) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sq" complement(9318..9442) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J or Alu-S" 9486..9773 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J" 9996..10132 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sp" 10159..10255 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sq" 10271..10367 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J" complement(10455..10739) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sx" complement(10740..10860) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-J or Alu-S" complement(10866..11156) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sx" 11223..11501 /evidence=EXPERIMENTAL /map="X"
repeat_region	/rpt_family="Alu-J" 11638..13564 /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="CpG Island" complement(11884..12163) /evidence=EXPERIMENTAL /map="X"
repeat_unit	/rpt_family="Alu-Sb0 or Alu-Sb1" complement(12195..12482) /evidence=EXPERIMENTAL



```
FT /map="X"
FT /rpt_family="Alu-J or Alu-S"
FT complement (12668..12945)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J or Alu-S"
FT complement (18360..18648)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sb0"
FT complement (18984..19488)
FT /evidence=EXPERIMENTAL
FT /gene="FLN"
FT /map="X"
FT /number=48
FT complement (join(19301..19488,19814..20017,20101..20319,
FT 20483..20659,21361..21493,22033..22148,22336..22473,22633.
FT 22899,23005..23127,23224..23376,23453..23656,23748..23909,
FT 24006..24179,24367..24495,24603..24743,24833..24935,25068.
FT 25163,25277..25524,27701..27724,27884..28073,28649..28805,
FT 28895..29018,29434..29604,29696..29856,29934..30096,30182.
FT 30355,30440..31037,31758..32020,32120..32237,32429..32598,
FT 32692..32782,32868..33028,33111..33234,34472..34615,34709.
FT 34822,34976..35169,35271..35407,35586..35709,35798..35935,
FT 36473..36673,36757..36919,37011..37088,37181..37299,37846.
FT 37993,38090..38187,38291..38539,41319..41691))
FT /evidence=EXPERIMENTAL
FT /gene="FLN"
FT /map="X"
FT /codon_start=1
FT /functions="binds actin"
FT /product="filamin"
FT /db_xref="PDB:1203969"
FT /translation="MSSSHRSRAGQSAAGAPGGVTRDAEMPATEKDIAEDAPWKI
Q
...
Note: remainder of annotations omitted.

Query Match 14.9%; Score 233; DB 10; Length 219447;
Best Local Similarity 90.0%; Pred. No. 3,45e-101;
Matches 262; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Db 7120 ttttttctgagatggatttccttcatttcgcagagctggagtgcaatggcacaatct 7179
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 969 TTTTTCGATGATGAGTTCGCTTTGCTGCCCCAGGCTGAGTGCAATGCCGATCT 1028
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 7180 tgccaccgcaacctcgtcccggtttgagcgattctcctgctcagcctcccgag 7239
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 1029 TGCTCACCATAACCTCCGCGCTCCAGGTCGAGCAATTCCTCGCTTAGCGCTGAG 1088
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 7240 taactggatcacagcattgacacattgctgcttattttatttttagtagagac 7299
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 1089 TAGCTGGATTACAGCGCTGCGGCACTATGCTGACTAATTTTCTAGTTAGTAGAGAC 1148
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 7300 aggttttccatgtggctcagctggctcctcaactcgaactcaggtgatctgccgcg 7359
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Qy 1149 GGGGTTTCTCCATGTTGGTCAGGCTGCTCCAACTCCTGACCTCAGGTGATCTGCCCG 1208
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 7360 ctgagctcccaatgtctgaattacaggtgtgagccaccacgcccgct 7410
```

```
Qy 1209 CTCAGCTCCCAAGTCTGGAATTACGGCTGAGCCACCCCTGGCT 1259
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
RESULT 13
ID HSFING6PD standard; DNA; PRI; 219447 BP.
AC L44140;
DT 29-FEB-1996 (Rel. 47, Last updated, Version 2)
DE Homo sapiens chromosome X region from filamin (FLN) gene to
DE glucose-6-phosphate dehydrogenase (G6PD) gene, complete cds's.
KW 1A gene; 2.19 gene; ABP-280 gene; actin-binding protein;
KW DNase I-like protein gene; DNLI1 gene; EMD gene; emerlin;
KW emery-dreyfuss syndrome; filamin; FLN gene; G4.5 gene;
KW G6PD gene; gdi gene; GdX gene; glucose-6-phosphate dehydrogenase;
KW P3 gene; QM gene; STA gene; XAP-1 gene; XAP-2 gene; XAP-4 gene;
KW XAP-5 gene; XAP-7 gene.
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
RN [1]
RP 1-219447
RA Chen E.Y., Zollo M., Mazzarella R.A., Ciccodicola A., Chen C.N.,
RA Zuo L., Helner C., Burrough F.W., Ripetto M., Schlessinger D.,
RA D'Urso M.;
RT "Thirteen known and six candidate genes in 219.4Kb of high GC DNA
RT between the human RCP/GCP and G6PD loci";
RL Unpublished.
CC Submitted by: Ellison Chen, Advanced Center for Genetic Technology,
CC Applied Biosystems Division of Perlin Elmer Corp., 850 Lincoln
CC Center Drive, Foster City, CA 94004 USA and David Schlessinger,
CC Department of Molecular Microbiology and Center for Genetics in
CC Medicine Washington University School of Medicine, St. Louis MO
CC 63110 USA e-mail: ellison@geneseq.apldbio.com and
CC david@genetics.wustl.edu Note: Gene predictions were accomplished
CC with runs of Grail versions 1.1 and 1.2, coupled with fasta and
CC blastx comparisons to genbank & non-redundant peptide libraries.
CC Repeat analysis was accomplished via censor. NCBI gi: 1203968
FH Key Location/Qualifiers
FT source 1..219447
FT /organism="Homo sapiens"
FT /sequenced_mol="DNA"
FT repeat_unit 4..223
FT /evidence=EXPERIMENTAL
FT /map="X"
FT repeat_unit /rpt_family="Alu-Sp or Alu-Sq"
FT 226..503
FT /evidence=EXPERIMENTAL
FT /map="X"
FT repeat_unit /rpt_family="Alu-J or Alu-S"
FT 641..931
FT /evidence=EXPERIMENTAL
FT /map="X"
FT repeat_unit /rpt_family="Alu-J or Alu-S"
FT 1023..1313
FT /evidence=EXPERIMENTAL
FT /map="X"
FT repeat_unit /rpt_family="Alu-Sx"
FT complement(1446..1536)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT repeat_unit /rpt_family="Alu-J"
FT 3095..3426
```

Aug 27 08:17

US-08-231-565A-1.rge

35

```
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="CpG Island"
FT complement(3663..3952)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sb0 or Alu-Sbl"
FT 4912..5046
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J or Alu-S"
FT 5047..5333
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sb0 or Alu-Sbl"
FT 5346..5628
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J or Alu-S"
FT 5632..5805
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sq"
FT 6363..17608
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="11Kb repeat1"
FT 6369..6633
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sx"
FT complement(7121..7409)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sp"
FT complement(7358..7432)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sb0 or Alu-Sbl"
FT complement(7426..7701)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J"
FT 8306..8592
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sx"
FT complement(8863..9029)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J or Alu-S"
FT complement(9031..9317)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sq"
FT complement(9318..9442)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J or Alu-S"
FT 9486..9773
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J"
FT 9996..10132
```

Aug 27 08:17

US-08-231-565A-1.rge

36

```
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sp"
FT 10159..10255
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sq"
FT 10271..10367
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J"
FT complement(10455..10739)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sx"
FT complement(10740..10860)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J or Alu-S"
FT complement(10866..11156)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sx"
FT 11223..11501
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J"
FT 11638..13564
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="CpG Island"
FT complement(11884..12163)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sb0 or Alu-Sbl"
FT complement(12195..12482)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J or Alu-S"
FT complement(12668..12945)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-J or Alu-S"
FT complement(18360..18648)
FT /evidence=EXPERIMENTAL
FT /map="X"
FT /rpt_family="Alu-Sb0"
FT complement(18984..19488)
FT /evidence=EXPERIMENTAL
FT /gene="FLN"
FT /map="X"
FT /number=48
FT complement(join(19301..19488,19814..20017,20101..20319,
20483..20659,21361..21493,22033..22148,22336..22473,22633,
22899,23005..23127,23224..23376,23453..23656,23748..23909,
24006..24179,24367..24495,24603..24743,24833..24935,25068,
25163,25277..25524,27701..27724,27884..28073,28649..28805,
28895..29018,29434..29604,29696..29856,29934..30096,30182,
30355,30440..31037,31758..32020,32120..32237,32429..32598,
32692..32782,32868..33028,33111..33234,34472..34615,34709,
```

241..312  
tRNA

few errors 88

```

CC The sequence from clone L174G8 has been finished in more than one
CC contig.
CC This sequence (L174G8C) is separated from the preceeding one
CC (L174G8B)
CC by a gap of about 1.2kb sized by restriction analysis.
CC The true right end of clone L174G8 is at 1846.
CC L174G8 is from cosmid library LA04NC01 constructed at the
CC Human Genome Center, Los Alamos National Laboratory, NM 87545
CC under the auspices of the U.S. Department of Energy.
CC The library was constructed using flow-sorted human chromosome 4
CC from a Hamster-Human hybrid cell line (UV20HL21-27) containing
CC human chromosomes 4, 8 and 21.
CC VECTOR: sCos1
CC L174G8 is contained in a clone contig spanning ~2Mb which is
CC described in Baxendale et al, Nature Genetics 4 ( 1993 ) 181-
CC 186. See also Myers et al, Cytogenet Cell Genet. 66 ( 1994 )
CC 218-230.
FH Key Location/Qualifiers
FT source 1..1846
FT /organism="Homo sapiens"
FT /chromosomes="4"
FT /map="4p16.3"
FT /clone="L174G8"
FT /clone_lib="LA04NC01"
FT /cell_line="UV20HL21-27"
FT /note="969..1203
FT repeat_region /note="Alu repeat: matches 60..308 of consensus"
FT /partial
FT repeat_region 1255..1546
FT /note="Alu repeat: matches 308..1 of consensus"
FT /partial
FT repeat_region 2..272
FT /note="Alu repeat: matches 287..1 of consensus"
FT /partial
SQ Sequence 1846 BP; 530 A; 440 C; 371 G; 505 T; 0 other;
Query Match 14.8%; Score 230; DB 10; Length 1846;
Best Local Similarity 89.9%; Pred. No. 1.39e-99;
Matches 259; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
Db 1254 tcttttttttgagatggagttgtctctgttgcacagctggagtgcaatggcacgat 1313
|||
967 TCTTTTTTTTGAGATGGAGTTGCGTTTCTTGGCCAGCGTGCAATGGCGCAT 1026
Db 1314 ctcggtcactgcaacctccgctccgggtgtaagccattctctgcctcagcctctg 1373
|||
1027 CTGGCTCACCATACCTCCGCCCTCCAGGTTCAACGAATTCCTCGCTTAGCCTCTG 1086
Db 1374 agtgcgtggtattacagcgccacacaccccggtcaattttgtatttttagtagag 1433
|||
1087 AGTACCTGGCATTACAGCGCGTGCGCACATATGCTCTACTAATTTGTAGTTAGTAGAG 1146
Db 1434 acagggtttctccatgttgcaggctgttctgaactccgacctcaggatgatccgct 1493
|||
1147 ACGGGTTTCTCCATTTGCTGAGCTGGTCTCAAACTCCTGACCTCAGGTATCTGCC 1206
Db 1494 gctcagcctcccaagtgtggattacaggcatgagccacgcgcc 1541
|||
1207 GCTCAGCCTCCCAAGTGTGGAATTACAGGCTGAGCCACACGCC 1254

```

Search completed: Tue Aug 27 08:29:27 1996

Job time : 1307 secs.

Result	No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
	1	1559	100.0	1559	17	T02714	MART-1 melanoma anti	0.00e+00
	2	633	40.6	760	17	T11550	Tumour rejection anti	0.00e+00
	3	350	22.5	354	12	Q76370	Tumour rejection anti	2.45e-1707
	4	304	19.5	13585	17	T11549	Tumour rejection anti	3.98e-1451
c	5	232	14.9	6905	15	Q92779	Human thymopoietin co	6.55e-108
	6	215	13.8	370	8	Q59541	Human brain Expressed	1.04e-96
	7	215	13.8	387	8	Q59806	Human brain Expressed	1.04e-96
	8	212	13.6	8174	2	T13332	GDP-Fuc:beta-D-galact	4.32e-95
	9	212	13.6	8174	9	Q56908	DNA encoding a glycos	4.32e-95

RESULT	1
ID	T02714 standard; cDNA; 1559 BP.
AC	T02714,
DDT	20-APR-1996 (first entry)
DE	MART-1 melanoma antigen.
KW	MART-1; melanoma antigen recognised by T-cells; melanoma;
KW	metastatic melanoma; tumour-associated antigen; immunogen;
KW	diagnosis; prognosis; prophylaxis; therapy; vaccine; ds.
SOS	Mammalian.
Key	Location/Qualifiers
FT	Key 54..410
CDS	/*tag= a
PPN	W09529193-A2.
PPD	02-NOV-1995.
PPF	21-APR-1995; U05063.
PPR	22-APR-1994; US-231565.
PPR	05-APR-1995; US-417174.
PPI	(USSH ) US SEC DEPT HEALTH.
PPI	Kawakami Y, Rosenberg SA;
PPI	MP1; 95-382963/49.
PPI	p-PSDB; R84212.
PPI	DNA encoding melanoma antigens recognised by T-lymphocytes - also
PPI	vectors, host cells and antibodies, used to detect, treat and
PPI	immunise animal against melanoma.

PS Claim 2; Page 115-116; 184pp; English.

CC The nucleic acid encodes a melanoma antigen (MART-1) which is  
CC recognized by T-lymphocytes. It is used for recombinant protein  
CC production, preferably using a baculo virus vector for expression in  
CC insect cell cultures. MART-1 protein is a source of immunogenic  
CC peptides (see R84196 for peptide M9-2) which are optionally modified  
CC (see R84783-R84800) and used in medicaments for the treatment or  
CC prevention (by immunization) of melanoma. Antibodies against MART-1  
CC and its immunogenic peptides may be used in the detection and  
CC isolation of MART-1 from a sample, the detection of which is  
CC indicative of a disease state (melanoma or metastatic melanoma).  
SQ Sequence 1559 BP; 470 A; 330 C; 324 G; 435 T;

Query Match 100.0%; Score 1559; DB 17; Length 1559;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 1559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db 1 agcagacagagactctcatiaagaagtgctcctgtgccctgacctgacctacaagatgccaa 60
Qy 1 AGCAGACAGAGACTCTCATTAAGGAAGGTGCTCTGTGCCCTGCACCTACAAGATGCCAA 60

Db 61 gagaagatgctcaactcatctatggttaaccccaagaaggggcaaggccactcttacacca 120
Qy 61 GAGAGATGCTCACTTCATCTATGCTTACCCCAAGAGGGGCGACGCCACTCTTACACCA 120

Db 121 cggctgaagagcgctggagtcggcatcctgacagtgatcctggagtccttaactgtctca 180
Qy 121 CGGCTGAAGAGCGGCTGGGATCGGCATCCTTGACATGATCTCGGAGTCTTACTGCTCA 180

Db 181 tcggctgttggtatgtagaagacaaatggtatcacagagccttgatgataaaagtcttc 240
Qy 181 TCGGCTGTGTGATATGTAGAACAGCAAAATGGATACAGAGCCCTTGATGAATAAAGTCTTC 240

Db 241 atgttgcaactcaatgtccttaacagaagatgcccaagaaggggttgatcatcagg 300
Qy 241 ATGTGGCACTCAATGTGCTTAAACAAGAGATGCCACAGAGGGTGTGATCATCGGG 300

Db 301 acagcaagtgtctcttcaagagaaaaactgtgaacctgtggttcccaatgtccacctg 360
Qy 301 ACAGCAAGTGTCTCTTCAAGAGAAAATCTGAACTGTGTGCTGCCAATGCTCCACCTG 360

Db 361 cttatgagaaactctctgagaacagtcaccaccacttattcacccttaagagccagca 420
Qy 361 CTTATGAGAAACTCTCTGAGAACAGTCGACCAACCCCTTATTTCACCTTAAGAGCCAGCA 420

Db 421 gacacctgagacatgctgaatatttctctcacacttttgcttgaatttaacagaca 480
Qy 421 GACACCTGAGACATGCTGAAATATTTCCTCTCACACTTTGCTTGAATTAATACAGACA 480

Db 481 tctaattgtctccttggaaatggtgtagaaaaatgcaagccatctctcaataaagtca 540
Qy 481 TCTAATGTCTCCTTTGGAATGTTGAGAAAATGCAAGCCATCTCTAATAATAAGTCA 540

Db 541 gtgttaaatatttagtggtccgtagcagtaactaatcatgtgaggaatgatgagaat 600
Qy 541 GTGTAAATATTTAGTGTCCGTAGCAGTACTTAATCATGTGAGGAATGATGAGAAAT 600

Db 601 attaaatggaaaactccatcaataatgttgcaatgcatgatactatctgtgccagag 660
Qy 601 ATTAATTTGGGAAAATCTCCATCAATAAATGTTGCAATGCATCATCTGTGCCAGAG 660

Db 661 gtaatgttaaatccatggtgttatttctgagagacagaattcaagtgggtattctg 720
Qy 661 GTAATGTAGTAAATCCATGTTGTTATTTTCTGAGACAGAGAAATTCAGTGGGTATTCTG 720
```

```
Db 721 ggccatccaattctcttacttgaattggctaaatacaaaactagtcagggttttca 780
Qy 721 GGGCATCCAAATTTCTTTTACTTTGAATTTGGCTAATAACAAACTAGTCAGGTTTTCGA 780

Db 781 acctgacgcagatgaactgtacacagaatgttccagttactatggagtgctcacaaagg 840
Qy 781 ACCTTGCCGACATGAACGTGTACACAGAAATTTGTTCCAGTACTATGCGAGTGCTCAAAAG 840

Db 841 atactttacaggttaagcaaaaggttgactgacctatttatctgatcaagaacatgtc 900
Qy 841 ATACTTTACAGGTTAAGCAAAAGGTTGACTGCGCTATTATCTGATCAAGACATGTC 900

Db 901 agcaatgtctcttctgtcttaaaattctattatacaataatatttgaagatcc 960
Qy 901 AGCANGTCTCTTTGCTCTTAAATTTCTATTACTACAATAATATTGTAAGATCC 960

Db 961 tatagctcttttttttgagatggagtttcgcttttggccaggctggagtgcaatgg 1020
Qy 961 TATAGTCTTTTTTTTTCAGATGGAGTTTCGCTTTTGTGCCCAGGCTGGAGTGCAATGG 1020

Db 1021 cggatcttgcttccataaactccgctccaggctccagggttcaaggcaattctctgcttagc 1080
Qy 1021 CGGATCTTGGCTCAGCATTAACCTCCGCTCCAGGTTCAAGCAATTTCTCTGCTTACG 1080

Db 1081 ctcctgagtagctgggattacagcgctgcccactatgcttgactaattttgtagtctta 1140
Qy 1081 CTCCTGAGTAGCTGGGATTCAGGCGTGGGCCACTATGCTGACTAATTTTGTAGTTTAA 1140

Db 1141 gttagagcggggtttctccatgttggcaggtgtgtctcaaaactcctgacctcagggtgat 1200
Qy 1141 GTAGACGCGGGGTTTCTCCATGTTGGTCAAGGCTGGTCTCAAACTCTCGACCTCAGGTA 1200

Db 1201 ctgcccgcctcagcctcccaaatgctggaattacaggggtgagcccaacagcctggctg 1260
Qy 1201 CTGCGCGCCTCAGCCTCCCAAGTGTGGAATTCAGGGGTGAGCCACCAAGCCTGGCTG 1260

Db 1261 gatcctatatcttagtgaagacataaacgcagtcctaatacttaacttcaactcaaggctca 1320
Qy 1261 GATCCTATATCTTAGTGAAGACATAAACGCAGTCTAATTACATTTTCACTTCAAGGCTCA 1320

Db 1321 atgctattctaactaatgacaagtattttctactaaaccagaaatgtgtagaaggattta 1380
Qy 1321 ATGCTATTCTAATTAATGACAAGTATTTTCTACTTAACCAAGAAATGGTAGAAGATTAA 1380

Db 1381 aataagtaaaagtactatgtactgccttagtgctgactgcctttaaagt 1440
Qy 1381 AATAAGTAAAAAGCTACTATGTACTGCGCTTAGTGCTGATCCCTGTGTACTGCGCTTAAATGT 1440

Db 1441 acctatggcaatttagtctctgttcccaaatccctctccacaagaatgtgcagaaga 1500
Qy 1441 ACCTATGGCAATTTAGTCTCTTGGGTTCCCAATCCCTCTCAACAAGATGTGCAGAAGA 1500

Db 1501 aatcataaagatcagagatcttgaaataaaaaaataaaaaaataaaaaa 1559
Qy 1501 AATCATAAAGATCAGAGATCTCGAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1559
```

## RESULT 2

ID T11550 standard; cDNA; 760 BP.  
AC T11550;  
DT 18-APR-1996 (first entry)  
DE Tumour rejection antigen precursor coding sequence.  
KW Tumour rejection antigen precursor; TRAP; TRA; melanoma; cancer;





Db 121 ctctccatcgctgtgtgtatttagagacgaatggatagagccttgatgataaa 180  
 |||||  
 Qy 174 CTCTCATCGCGCTGTGGTATTGTACAGACGAATGGATACAGCGCTTGATGATAA 233  
 |||||  
 Db 181 agtcttcattgtgtgcaactcaatgtgcttcaagaagatgccacaagaagggtttgat 240  
 |||||  
 Qy 234 AGTCTTCATTGTGGCACTCAATGTGGCTTTAACAAGAGATGCCACAAGAGGGTTGTAT 293  
 |||||  
 Db 241 catcgagagacgaatgtctcttcaagagaaaactgtgaacctgtgttcccaatgct 300  
 |||||  
 Qy 294 CATCGGAGACGAAGTGTCTCTTCAAGAGAAAAGTGTGAACCTGTGTTCCCAATGCT 353  
 |||||  
 Db 301 ccactgttatgagaaactctctgcagaaactgtgcagaaactcaccaccacttattcaact 354  
 |||||  
 Qy 354 CCACCTGCTTATGAGAACTCTCTGCAGACAGACGTCCACCACCACCTATTACCT 407  
 |||||

## RESULT 4

ID T11549 standard; DNA; 13585 BP.  
 AC T11549;  
 DE Tumour rejection antigen precursor coding sequence.  
 KW Tumour rejection antigen precursor; TRAP; TRA; melanoma; cancer;  
 KW tumour; treatment; detection; vaccine; HLA-A2; adoptive transfer;  
 KW T cell; T lymphocyte; human leukocyte antigen; ds.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT misc feature 2685  
 FT /\*tag= a  
 FT /note= "Unidentified nucleotide."  
 FT misc feature 9422..9456  
 FT /\*tag= b  
 FT /note= "This region is 4.7-5.3 kilobases in length  
 but the sequence of this region has not been  
 deduced."  
 FT misc feature 11539  
 FT /\*tag= c  
 FT /note= "Unidentified nucleotide."  
 PN W09601557-A1.  
 PD 25-JAN-1996.  
 PF 27-JUN-1995; U08153.  
 PR 08-JUL-1994; US-212351.  
 PR 10-JAN-1995; US-370319.  
 PA (LUDW-) LUDWIG INST CANCER RES.  
 PI Boon-falleur T, Brichard V, Coulie P, De PLAEN E;  
 PI Traversari C, Van PEL A, Wolfel T;  
 DR WPI; 96-097390/10.  
 PT Genomic DNA encoding a tumour rejection antigen precursor -  
 processed to antigen presented by HLA-A2, useful for treating or  
 diagnosing melanoma  
 PT  
 PS Claim 3; Page 18-22; 41pp; English.  
 CC The tumour rejection antigen precursor (TRAP) is processed to a  
 tumour rejection antigen (TRA) presented by HLA-A2 molecules. TRA  
 is used to generate cytotoxic T lymphocytes for treating cancer (esp.  
 melanoma). It can also be used to raise specific antibodies, and  
 when complexed with HLA-A2, it can be used to produce vaccines.  
 CC Cytotoxic T lymphocytes so generated can be used in adoptive  
 CC transfer or generated, or they can be generated in vivo by using a  
 CC vector containing the appropriate gene or using TRA or TRAP together  
 CC with an adjuvant that facilitates entry into HLA-A2 presenting cells.  
 CC Diagnostic methods involving the detection of expression of TRAP  
 CC can be used in the detection of cancers.  
 SQ Sequence 13585 BP; 3827 A; 2859 C; 2968 G; 3894 T;

Query Match 19.5%; Score 304; DB 17; Length 13585;  
 Best Local Similarity 100.0%; Pred. No. 3.98e-145;  
 Matches 304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 13274 gttcccaatgctccacctgcttatgagaaactctctgcagaaactcaccaccacttat 13333  
 |||||  
 Qy 342 GTTCCCAATGCTCCACCTGCTTATGAGAACTCTCTGCAGACAGTCACCACCCTTAT 401  
 |||||  
 Db 13334 tcaccttaagagccagcagacacccctgagacatgctgaaattattctctcacactttg 13393  
 |||||  
 Qy 402 TCACCTTAAGAGCCAGCCAGACACCTGACACATCTGTAATATTTCTCTCACACTTTG 461  
 |||||  
 Db 13394 cttgaatttaatacagacaactaatgtctctctttgggaatggtgtaggaataatgcaagc 13453  
 |||||  
 Qy 462 CTTGAATTAATATACAGACATCTAATGTCTCTCTTGGAAATGCTAGGAAATCCAAGC 521  
 |||||  
 Db 13454 catctctaataaagtcagtggttaaaatttagtgctcgctagctagctactaatcatg 13513  
 |||||  
 Qy 522 CATCTCTAATAAAGTCAGTGTAAATTTTAGTGGCTAGCAGTACTAATCATG 581  
 |||||  
 Db 13514 tgagaaatgatgagaatattaaattgggaaaaactccatcaataaattgtgcaatgcat 13573  
 |||||  
 Qy 582 TGAGGAATGATGAGAAATATTAAATTTGGGAAATCTCCATCAATAATGTTCCAAATGCAT 641  
 |||||  
 Db 13574 gata 13577  
 |||||  
 Qy 642 GATA 645

## RESULT 5

ID Q92779 standard; DNA; 6905 BP.  
 AC Q92779;  
 DT 09-DEC-1995 (first entry)  
 DE Human thymopoietin continuous gene fragment.  
 KW Thymopoietin; diagnostic; therapeutic; gene therapy; ss.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT exon 39..11956-2843  
 FT /\*tag= a  
 FT /note= "exon 4"  
 FT exon 4691..4788  
 FT /\*tag= b  
 FT /note= "exon 5"  
 PN W09517205-A1.  
 PD 29-JUN-1995.  
 PF 13-DEC-1994; U14356.  
 PR 21-DEC-1993; US-171382.  
 PA (IMMU-) IMMUNOBIOLOGY RES INST INC.  
 PI Culler MD, Goldstein G, Harris CA, Setcavage DR;  
 PI Shenbagamurthi P, Siekierka JJ, Talle MA;  
 DR WPI; 95-240474/31.  
 PT Polynucleotide(s) encoding human thymopoietin proteins - used to  
 develop prods. for diagnosis and therapy involving immune or nervous  
 system conditions.  
 PS Disclosure; Fig 9a-9g; 85pp; English.  
 CC The sequence represents a fragment of the human thymopoietin gene,  
 CC including exons 4 and 5. DNA encoding human thymopoietin, from  
 CC a cDNA library prepared from human thymus RNA, may be expressed in  
 CC a host cell e.g. bacterium (preferably Escherichia coli), fungus,  
 CC insect or mammalian cell for production of recombinant thymopoietin.  
 CC Thymopoietin has a regulatory effect on the mammalian immune system,  
 CC and can be used for treating/modulating an immune or nervous system  
 CC condition, and for the treatment of chronic infection, autoimmune  
 CC disorders and certain affective psychiatric or neurological





QY	1106	GTGCGCCACTATGCGCTGACTAAATTTTGTAGTTTACTAGACAGGGGTTTCTCCATGTTG	1165
Db	180	gtcagcgtggtcttgaactcctcgaactcagggtatccacactgcctcggcctcccaaatg	239
QY	1166	GTACGGCTGGTCTCTAAACTCTGACCTCAGSGTAGTCCCGGCTCAGCGTCCCAAGTG	1225
Db	240	ctggattacaggcataagccactggcc	268
QY	1226	CTTGAATACAGCGCTGAGCCACACGGC	1254

RESULT	8	
ID	Q13332	standard; DNA; 8174 BP.
AC	Q13332;	
DT	07-NOV-1991	(first entry)
DE	GDP-Fuc:[beta-D-galactoside alpha(1,2)-fucosyltransferase gene.	
KE	Glycosyltransferase.	
OS	Homo sapiens.	
FH	Key	Location/Qualifiers
FT	CDS	4686..5783
FT	/*tag= a	
PN	W09112340-A.	
PD	22-AUG-1991.	
PF	14-FEB-1991;	U00899.
PR	14-FEB-1990;	US-480133.
PR	14-FEB-1990;	US-479858.
PR	12-DEC-1990;	US-627621.
PA	(UNMI ) UNIV OF MICHIGAN.	
PI	Lowe JB; .	
DR	WPI; 91-267151/36.	
DR	P-RSDB; R13751.	
PT	Isolation of gene conveying post-translational characteristic -	
PT	e.g. the presence of soluble or membrane bound oligo or	
PT	polysaccharide or glycosyltransferase.	
PS	Disclosure; Fig 3; 155pp; English.	
CC	The DNA encodes a protein sequence capable of functioning as a	
CC	GDP-Fuc:[beta-D-Gal alpha(1,2)- fucosyltransferase. The sequence	
CC	coded by nucleotides 4782 - 5780 represents the functional protein.	
CC	The enzyme produced by the DNA sequence can be used in enzymatic	
CC	fucosylation of chain-terminating galactose residues on lactose-	
CC	amine or neolacto type beta-D-galactoside to alpha-2-L-fucose	
CC	residues. See also Q13330-Q13333.	
!SQ	Sequence 8174 BP; 1628 A; 2229 C; 2322 G; 1995 T;	

Query Match	13.6%	Score 212;	DB 2;	Length 8174;
Best Local Similarity	86.6%	Pred. No. 4,32e-95;		
Matches 251;	Conservative 0;	Mismatches 39;	Indels 0;	Gaps 0;
Db	4007	tttttttttgagacgagtggttcaactcttgttcccaaggctggagtgtaatggcatgatct	4066	
Qy	969	TTTTTTTTTGATGGAGTTTGGCTTTTGTGCCAGCTCGAGTCCAATGGCCGATCT	1028	
Db	4067	cagctcacccgaactccgcgcccccgggttcaggcgattctcctgcctcagcctccccag	4126	
Qy	1029	TGGCTCAGCATTAACCTCCGCCCTCCGAGTTTCAAGCAATTCTCGCTTAGCTCTCGAG	1088	
Db	4127	tgctggattacagcgcatgcgcacacatgcccggtcaattttgtattttaagtatagac	4186	
Qy	1089	TAGCTGGGATTACAGCGCTCGCCCATATCGCTGACTAAATTTTGTAGTTTGTAGTAGAC	1148	
Db	4187	agggtttctccaagtttggtcaggctggtctgaactccccaaacctcagtgatccacccac	4246	
Qy	1149	GGGTTTTTCATTTGTTGTCAGGTTGCTTCAACTCTTGACCTAGGTGATGTCGCCGC	1208	

```
Db 4247 ctggtcccccacaaagtctggattacaggttgagccaccgcgcctggc 4296
||| |||||||||||||||||||||
0w 1209 CTCACGCTCCCAAGTGTCTGGATTACAGCGGTGAGGCCACACGCCTGGC 1258
```

RESULT	9
ID	Q56908 standard; DNA; 8174 BP.
AC	Q56908;
DT	26-JUL-1994 (first entry)
DE	DNA encoding a glycosyltransferase.
KW	Glycosyltransferase; fucosyltransferase; GDP-Fuc; in vitro; cell;
KW	surface; oligosaccharide; ss.
OS	Homo sapiens.
Key	Location/Qualifiers
FT	CDS
FT	/tag= a
PN	M09402616-A.
PD	03-FEB-1994.
PF	20-JUL-1993; U06703.
PR	20-JUL-1992; US-914281.
PA	(UNMI ) UNIV MICHIGAN.
PI	Lowe JB,
DR	WPI; 94-048874/06.
P-	PSDB; R45936.
PT	DNA fragment encoding a glycosyltransferase - can be used for in
PT	vivo reactions to modify cell surface oligosaccharide(s) e.g.
PT	blood g. determinants, to protect against transplant rejection
PS	Disclosure; Fig 3; 24pp; English.
CC	The sequence is that encoding human glycosyl transferase. The en
CC	produced by the DNA may be non glycosylated. This prevents prena
CC	loss of enzyme activity. It can also be used in in vitro reactio
CC	modify cell surface oligosaccharide mols. e.g. blood group determ
CC	See also Q56905-12.
SQ	Sequence 8174 BP; 1628 A; 2228 C; 2322 G; 1996 T;

Query Match	13.6%;	Score 212;	DB 9;	Length 8174;
Best Local Similarity	86.6%;	Pred. No. 4,32e-95;		
Matches 251;	Conservative	0;	Mismatches 39;	Indels 0;
Gaps 0;				
Db	4007	tttttttggagacggagtgttcaactctgttggccaggctggagtgaatggcatgatct	4066	
Qy	969	TTTTTTTTTGAGATGGAATTCGGTTCCTTTGCCCCAGCTGCAATGGCGGCATCT	1028	
Db	4067	cagctcaacgcaactccgcccgcgggttcaggcgattctcgtcctcagcctcccgag	4126	
Qy	1029	TGGCTGCACATACCTCGCGCTCCAGGTTCAAGCAATTCCTCGCCTTAGCCTCCTGAG	1088	
Db	4127	tgcctgggattacaggcatgcgcaccatgcccggtctaatttgtatttaagttagagac	4186	
Qy	1089	TAGCTGGGATATACAGCGCTGGCGCACTATGCTCTGACTAATTTTGTACTTTTGTAGTAGAC	1148	
Db	4187	agggtttctccagtttgtcaggctggtctcgaactcccaacctcagggtgatccaccac	4246	
Qy	1149	GGGGTTTCTCCATTTGGCTAGGCTGCTCAAACTCCTGAGCTCAGGTCACTCGCCGC	1208	
Db	4247	cttgcgtcccaagtgcctgggattacaggtgtgagccacgcgcctcgcg	4296	
Qy	1209	CTCAGCCTCCCAAGTGTGCAATTAACGGGTGAGCCACACAGCCCTGGC	1258	

RESULT 10  
ID N81564 standard; DNA; 11715 BP.  
AC N81564;  
DT 05-DEC-1990 (first entry)

DE Genomic sequence encoding protein C.  
KW Human protein C; blood coagulation disorders; ss.  
FH Key Location/Qualifiers  
FT exon 2131..2200  
FT /\*tag= a 3464..3630  
FT /label=exon I  
FT exon  
FT /\*tag= b 5093..5117  
FT /label=exon II  
FT exon  
FT /\*tag= c 5212..5349  
FT /label=exon III  
FT exon  
FT /\*tag= d 5452..5586  
FT /label=exon IV  
FT exon  
FT /\*tag= e 8255..8397  
FT /label=exon VI  
FT exon  
FT /\*tag= f 9270..9387  
FT /label=exon VII  
FT exon  
FT /\*tag= g 10517..12006  
FT /label=exon VIII  
FT exon  
FT /\*tag= h  
FT /label=exon IX  
PN EP-266190-A.  
PD 04-MAY-1988.  
PE 28-OCT-1987; 309528.  
PR 29-OCT-1986; US-924462.  
PA (ZYMO-) Zymogenetics Inc.  
PI Foster DC, Murray MJ, Berkner KL.;  
DR WPI; 88-121259/18.  
DR P-PSDB; P81205.  
PT Protein C DNA coding sequence and expression vector for prodn. -  
PT used for treating blood coagulation disorders.  
PS Disclosure; p; English.  
CC In the construction of the full length protein C gene this sequence  
CC the introns removed then is joined to a cDNA clone. The desired  
CC genomic:cDNA is then generated by looping out of unwanted sequences  
CC using oligonucleotide-directed deletion mutagenesis. The protein  
CC produced upon transformation of mammalian host cells, contg. the  
CC recombinant DNA, has substantially the same biological activity  
CC as natural protein C and is hence useful in the treatment of blood  
CC coagulation disorders. See also N81563.  
SQ Sequence 11715 BP; 2443 A; 3292 G; 3375 G; 2605 T;

Query Match 13.6%; Score 212; DB 1; Length 11715;  
Best Local Similarity 87.5%; Pred. No. 4.32e-95;  
Matches 253; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

Db 6376 tcttttttttgatgagtttcaactcttctgtccagctgagtgagtgcaatgacgtga 6435  
|||||  
QY 967 TCTTTTGTGATGAGTTTCGCTTTGTTG-CCAGCTGGAGTCAATGGCGCA 1025  
|||||  
Db 6436 tctcagctcaccacaacctccgctctggattcaagcgattctctcgcgagctccc 6495  
|||  
QY 1026 TCTTGCTCACCAATACCTCGCGCTCCAGGTTCAAGCAATTCCTCGCTTAGCCTCT 1085  
|||||  
Db 6496 gadtgctggattacagcagtcgcaccacgcagctaatttgtgttttagtaga 6555  
|||||  
QY 1086 GAGTAGCTGGATTACAGCGCTGGCCACTATGCTGACTAATTTCTAGTTTAGTAGA 1145  
|||||

Db 6556 gaagggtttctcgtgtgtgtcaagctggtcttgaaactcctgacctcagtgatccacc 6615  
|||  
QY 1146 GACGGGGTTTCTCATGTGCTCAGGCTGGTCTCAAACTCTGACCTCAGGTGATCTGCC 1205  
|||||  
Db 6616 tgcttgctctcctaaagtctggtgattacagggtgagccaccgcgc 6664  
|||||  
QY 1206 CGGCTCAGCTCCCAAGTGTCTGAATTACAGGGGTGAGCCACCAAGCC 1254  
|||||

RESULT 11  
ID N70102 standard; DNA; 11724 BP.  
AC N70102;  
DT 10-MAY-1991 (first entry)  
DE Complete genomic sequence of human Protein C.  
KW Human Protein C; anti-coagulant; thrombosis; serine protease; ss.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT exon 2131..2200  
FT /\*tag= a  
FT /number= 1  
FT /note= "encodes amino acids -42 to -20"  
FT intron 2201..3463  
FT /\*tag= b  
FT /number= 1  
FT exon 2464..3630  
FT /\*tag= c  
FT /number= 2  
FT /note= "encodes amino acids -19 to +37"  
FT intron 3631..5092  
FT /\*tag= d  
FT /number= 2  
FT exon 5093..5117  
FT /\*tag= e  
FT /number= 3  
FT /note= "encodes amino acids 38 to 45"  
FT intron 5118..5209  
FT /\*tag= f  
FT /number= 3  
FT exon 5210..5347  
FT /\*tag= g  
FT /number= 4  
FT /note= "encodes amino acids 46 to 91"  
FT intron 5348..5449  
FT /\*tag= h  
FT /number= 4  
FT exon 5450..5584  
FT /\*tag= i  
FT /number= 5  
FT /note= "encodes amino acids 92 to 136"  
FT intron 5585..8252  
FT /\*tag= j  
FT /number= 5  
FT exon 8253..8395  
FT /\*tag= k  
FT /number= 6  
FT /note= "encodes amino acids 137 to 184"  
FT intron 8396..9267  
FT /\*tag= l  
FT /number= 6  
FT exon 9268..9385  
FT /\*tag= m  
FT /number= 7  
FT /note= "encodes amino acids 185 to 223"  
FT intron 9386..10514

Aug 27 08:19

US-08-231-565A-1.mg

15

FT /\*tag= n  
 FT /number= 7  
 FT exon  
 FT /\*tag= o  
 FT /number= 8  
 FT /note= "encodes amino acids 224 to 419"  
 FT polyA\_signal 11151..11156  
 FT /\*tag= q  
 FT polyA\_site 11173  
 FT /\*tag= i  
 FT polyA\_signal 11380..11385  
 FT /\*tag= g  
 FT polyA\_site 11399  
 FT /\*tag= t  
 PN EP-215548-A.  
 PD 25-MAR-1987.  
 PF 26-JUN-1986; 304970.  
 PR 27-JUN-1985; US-749600.  
 PR 15-AUG-1985; US-766109.  
 PA (ZYMO-) ZYMOGENETICS INC.  
 PA (UNIV ) UNIV OF WASHINGTON.  
 PI Murray MJ, Berkner KL, Foster DC, Davie EW.  
 DR WP1; 87-081505/12.  
 DR P-PSDB; P70855.  
 PT Human protein C or activated protein C - prep. using expression  
 PT vector capable of integration in mammalian host cell DNA  
 PS Claim 2; Page 33; 52pp; English.  
 CC A lambda gt cDNA library was prepared from human liver mRNA. The  
 CC library was screened with iodine-125 labelled antibody to human  
 CC protein C. A cDNA which lacked the coding sequence for the pre-pro  
 CC peptide and the first 23 amino acids of protein C was isolated and  
 CC used as a hybridisation probe. The remainder of the coding sequence  
 CC was obtained from a human genomic DNA library in lambda Charon 4A.  
 CC Recombinantly produced protein C can be used to treat thrombotic  
 CC disorders such as venous thrombosis.  
 SQ Sequence 11724 BP; 2445 A; 3302 C; 3370 G; 2607 T;

Query Match 13.6%; Score 212; DB 2; Length 11724;  
 Best Local Similarity 87.5%; Pred. No. 4.32e-95;  
 Matches 253; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

Db 6374 tcttttttttgatgaggtttcactcttctgtccaggctggagtgcaatgacgtga 6433  
 |||||  
 Qy 967 TCTTTTTTTTGATGGAGTTTCGCTTTGTTG-CCAGGCTGGAGTCAATGGCGGA 1025  
 |||||  
 Db 6434 tctcagctcaccacaacctcgcctcctggatcaagcattctcctgcgcagctccc 6493  
 |||||  
 Qy 1026 TCTTGCTCACCATAACCTCCGCTCCAGGTTCAAGCAATTCCTCGCTTAGCCTCT 1085  
 |||||  
 Db 6494 gagtagctgggattacaggaatcgccaccacgcccagctaatgtgttttagtaga 6553  
 |||||  
 Qy 1086 GACTAGCTGGGATTACAGCGTGGCCACTATGCTGACTAAATTTTGTAGTTTAGTGA 1145  
 |||||  
 Db 6554 gaaggggttctcgtgtgtgtgcaagctgggtcttgactctcctgacctcaggtgatcaacc 6613  
 |||||  
 Qy 1146 GACGGGGTTCTCCATGTTGGTCAGCGTGGTCTCAAACTCTGACCTGAGTGTCTGCC 1205  
 |||||  
 Db 6614 tgcctggcctcctaaagtctgggattacaggcgtgagccaccgcgc 6662  
 |||||  
 Qy 1206 CGCTCAGCTCCCAAGTGTGTAATTACAGGGCTGAGCCACACGCC 1254  
 |||||

RESULT 12  
 ID Q85367 standard; cDNA; 4788 BP.

Aug 27 08:19

US-08-231-565A-1.mg

16

AC Q85367;  
 DT 29-AUG-1995 (first entry)  
 DE MIP-1-alpha.  
 KW Macrophage inflammatory protein 1-alpha; MIP-1-alpha;  
 KW heparanase; heparin; heparan sulfate; arthritis; restenosis;  
 KW cancer; wound healing; ss.  
 OS Homo sapiens.  
 PN W09504158-A.  
 PD 09-FEB-1995.  
 PF 26-JUL-1994; U08207.  
 PR 29-JUL-1993; US-099866.  
 PR 13-OCT-1993; US-136117.  
 PA (UPJO ) UPJOHN CO.  
 PI Hoogwerf AJ, Ledbetter SR;  
 DR WP1; 95-082239/11.  
 DR P-PSDB; R70791.  
 PT Screening for cpds. with anti-heparanase activity - by detecting  
 PT inhibition of heparin or heparan sulphate degradation, cancer.  
 PT potentially useful for treating arthritis, restenosis, cancer.  
 PS Disclosure; Page 44-47; 60pp; English.  
 CC Purified heparanases, prepared under reducing conditions and  
 CC activated with transglutaminase, are given in R70786-805. Most  
 CC are prepared by reverse transcription of mRNA from activated human  
 CC leukocytes, then cloning of the cDNA into pVL1392 baculovirus  
 CC vector, and expression in Sf9 cells in the presence of reduced  
 CC glutathione and dithiothreitol.  
 SQ Sequence 4788 BP; 1299 A; 1122 C; 1049 G; 1318 T;

Query Match 13.3%; Score 207; DB 14; Length 4788;  
 Best Local Similarity 87.5%; Pred. No. 2.16e-92;  
 Matches 259; Conservative 0; Mismatches 34; Indels 3; Gaps 3;

Db 1533 ttttttttgatgaggtttcactcttctgtccaggctggagtgcaatgacgtgatct 1592  
 |||||  
 Qy 969 TTTTTTTTGATGGAGTTTCGCTTTTGTGCGGCTGAGTCAAGCAATTCCTCGCTT-AGCCTCCTGA 1028  
 |||||  
 Db 1593 cagctcaccacagcccccactcctcctgggttcaagcattctcctgcctcagctccaga 1652  
 |||||  
 Qy 1029 TGCTCACCATAACCTCCGCTCCAGGTTCAAGCAATTCCTCGCTT-AGCCTCCTGA 1087  
 |||||  
 Db 1653 gtagctgggattacaggaatgcccaccacgctgactacttttttagtagtaga 1712  
 |||||  
 Qy 1088 GTAGCTGGGATTACAGCGTGGCCACTATGCTGACTAAATTTTGTAGTTTAGTAGAGA 1147  
 |||||  
 Db 1713 tggagtttctcttctgtgtcaggttgggtcctcaactcctgacctcaggtgatctga-g 1771  
 |||||  
 Qy 1148 CGGGGTTTTCGATGTTGGTCAGGCTGGTCTCAAACTCTGACCTCAGGTGATCTGCCG 1207  
 |||||  
 Db 1772 cctcggcctcc-aaagtgtgggattacaggtgtgagcgaccatgctggctgcat 1826  
 |||||  
 Qy 1208 CCTCAGCCTCCCAAGTCTGGAATTACAGCGCTGAGCCACCGCTGGCTGGAT 1263  
 |||||

RESULT 13  
 ID Q54222 standard; DNA; 11531 BP.  
 AC Q54222;  
 DT 27-JUN-1994 (first entry)  
 DE BSSL/CEL Gene.  
 KW BSSL; biologically functional bile salt stimulated lipase; CEL;  
 KW carboxyl ester lipase; hybridisation; milk protein; transgenic;  
 KW infant milk substitute; oral; specific; lipolysis; vitamin; bile;  
 KW lipid malabsorption; cystic fibrosis; chronic pancreatitis;  
 KW digestion; ss.  
 OS Homo sapiens.



Aug 27 08:19

US-08-231-565A-1.mg

19

Cp 1075 GCGAGGAGATTCTTGAACCTGGAGGCGGAGGTTATGCTGAGCCACAGATCGGCCATT 1016

Db 246 gcaactcagcctggcgacaga-gcgagactcagctca 283

|||||

Cp 1015 GCACTCCAGCCTGGCAACAAAGCGAACTCCATCTCA 977

# RESULT 15

ID Q59876 standard; cDNA; 257 BP.

AC Q59876;

DT 16-MAR-1994 (first entry)

DE Human brain Expressed Sequence Tag EST01565.

KW Gene transcription product; genetic markers; tagging; in vivo;

KW transcription; mapping; location; chromosomes; chromosomal; ss.

OS Homo sapiens.

PN W09316178-A.

PD 19-AUG-1993.

PF 12-FEB-1993; U01294.

PR 12-FEB-1992; US-837195.

PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.

PI Adams MD, Moreno RF, Venter CJ.

DR WPI; 93-272882/34.

PT Enriched oligonucleotides and corresp. sequences - used as

PT markers for human genes transcribed in-vivo, facilitate tagging

PT of most human genes

PS Example 4; Page 246; 500pp; English.

CC The Expressed Sequence Tag was isolated from a human brain cDNA

CC library as part of a large set of ESTs which can be used as markers

CC for human genes transcribed in vivo. They can be used to facilitate

CC tagging of most human genes, for mapping locations of expressed genes

CC on chromosomes, for individual or forensic identification, for mapping

CC locations of disease-associated genes, for identification of tissue

CC type, and for prepn. of antisense sequences, probes and constructs.

CC EST01565 has a "poor" coding probability as evaluated using the

CC coding-region prediction program CRM. See also Q59041-Q61440.

SQ Sequence 257 BP; 48 A; 75 C; 60 G; 70 T;

Query Match 13.0%; Score 203; DB 8; Length 257;

Best Local Similarity 89.5%; Pred. No. 3.10e-90;

Matches 231; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

Db 1 ttctccctcgttgccaggtgagtgcaatgctgctcancctcccgagtagctgggattacaggca 120

|||||

Qy 987 TTTGCGTTTGTGGCCAGCGCTGGAGTGGCGGATCTTGGCTCAGCATACCTCC 1046

Db 61 gctccaccagctcagcaattctcctgctcancctcccgagtagctgggattacaggca 120

|||||

Qy 1047 GCTCCGAGGTTTCAGCAATTCCTCGCTTTCCTGCTAGTAGTGGATTACAGGGG 1106

Db 121 tgnccaccagctcagctcagcaatttnta-tttaagtagagagtggtttctccatgttg 179

|||||

Qy 1107 TGGCCACTATGCTGACTAATTTTGTAGTTAGTAGAGAGCGGGGTTTCTCCATGTGG 1166

Db 180 tcagtcgtggtctcaactcagctcaggtgagctggtccacctcgccctcccaaatgac 239

|||||

Qy 1167 TCAGGCTGGTCTCAAACTCCCTCAGCTCAGGTGATCTGCCGCCCTCAGCGCTCCCAAGTGC 1226

Db 240 tgggattacaggtgtgag 257

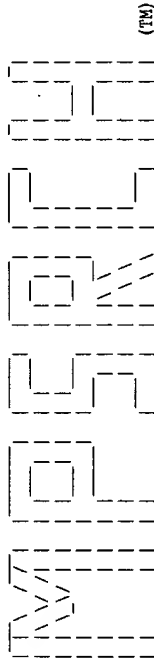
|||||

Qy 1227 TGAATTACAGCGCTGAG 1244

Search completed: Tue Aug 27 08:32:11 1996

Job time : 147 secs.

\*\*\*\*\*



(TM)

\*\*\*\*\*

Release 2.1D John F. Colline, Biocomputing Research Unit.  
 Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
 Distribution rights by IntelliGenetics, Inc.

MPsrch\_tpn n.a. - n.a. Smith-Waterman search, using a protein query  
 which has been backtranslated into n.a. using IUPAC symbols

Run on: Tue Aug 27 18:37:02 1996; MasPar time 260.36 Seconds  
 Tabular output not generated. 1033.428 Million cell updates/sec

Title: >US-08-231-565A-2  
 Description: (1-118) from US08231565A.pep  
 Perfect Score: 1420  
 N.A. Sequence: 1 ATCCGNNNGARGAYGNCAC.....SNCNCNCNCTAATWSNCCN 354  
 Comp: TACGKNCNCTCTCAGCT.....MNGGNGGNGATWSNCCN

Scoring table: TABLE bktranslate2  
 Gap 30

Nmatch STD : Dbase 0; Query 0

Searched: 270440 seqs, 380027776 bases x 2  
 Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: embl-new3  
 1: BCT 2: FUN 3: INV1 4: INV2 5: INV3 6: MAM 7: ORG 8: PLN  
 9: PRI1 10: PRI2 11: PRO 12: ROD 13: SYN 14: UNC 15: VRT  
 16: VIR1 17: VIR2

Database: genbank92  
 18: BCT1 19: BCT2 20: BCT3 21: BCT4 22: BCT5 23: BCT6 24: BCT7  
 25: BCT8 26: BCT9 27: INV1 28: INV2 29: INV3 30: INV4 31: INV5  
 32: INV6 33: MAM1 34: MAM2 35: PAT1 36: PAT2 37: PAT3 38: PHG  
 39: PLN1 40: PLN2 41: PLN3 42: PLN4 43: PLN5 44: PLN6 45: PLN7  
 46: PLN8 47: PRI1 48: PRI2 49: PRI3 50: PRI4 51: PRI5 52: PRI6  
 53: PRI7 54: PRI8 55: PRI9 56: PRI10 57: PRI11 58: PRI12  
 59: PRI13 60: ROD1 61: ROD2 62: ROD3 63: ROD4 64: ROD5 65: ROD6  
 66: ROD7 67: STR 68: SYN 69: UNA 70: VRL1 71: VRL2 72: VRL3  
 73: VRL4 74: VRL5 75: VRL6 76: VRL7 77: VRT1 78: VRT2 79: VRT3  
 genbank-new1  
 80: BCT 81: INV1 82: INV2 83: MAM 84: PHG 85: PLN 86: PRI  
 87: ROD 88: STR 89: SYN 90: UNA 91: VRL 92: VRT  
 u-embl45 92  
 93: part1

Database: Statistics: Mean 63.591; Variance 167.878; scale 0.379

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.	
1	1392	98.0	752	52	HSU06654	Human differentiation	1.18e-183	
2	1392	98.0	1524	52	HSU06452	Human melanoma antigen	1.18e-183	
3	172	12.1	12857	59	S77127	manganese superoxide	9.32e-04	
4	147	10.4	41591	27	ACU12386	Acanthamoeba castella	3.52e-01	
5	142	10.0	6004	52	HSS002G	H.sapiens SOD-2 gene	1.08e+00	
6	138	9.7	1467	36	A21548	probe for the identif	2.58e+00	
7	138	9.7	1467	22	LEPREL	Leptospira hardjo rep	2.58e+00	
8	137	9.6	267	50	HSGLNRS09	H.sapiens gene for gl	3.21e+00	
c	9	137	9.6	1384	28	CTBR21	Chironomus tentans Ba	3.21e+00
c	10	137	9.6	3347	77	CHKTGFA	Gallus gallus transfo	3.21e+00
c	11	137	9.6	7471	77	BRNOTCH	B.erio Notch mRNA.	3.21e+00
c	12	137	9.6	24537	28	CELF47F2	Caenorhabditis elegans	3.21e+00
13	133	9.4	1458	22	LBISLIKE	L.borgpetersenii IS-1	7.52e+00	
c	14	133	9.4	3287	22	HEUNIFABX	Herbaspirillum serop	7.52e+00
15	132	9.3	708	34	OATCVG6	O.aries rearranged T-	9.28e+00	
16	132	9.3	767	34	OATCCG4	O.aries rearranged T-	9.28e+00	
17	132	9.3	2208	34	RABRBAT	Oryctolagus cuniculus	9.28e+00	
18	132	9.3	2254	34	ACRRBATPRO	O. cuniculus mRNA for	9.28e+00	
19	132	9.3	3556	71	GVTVEF	Trichoplusia ni granu	9.28e+00	
20	132	9.3	3572	16	BAUGVEP	Pseudalattia unipuncta	9.28e+00	
21	132	9.3	3572	91	UGVEP	Pseudalattia unipuncta	9.28e+00	
22	132	9.3	3572	76	UGVEP	Pseudalattia unipuncta	9.28e+00	
23	132	9.3	3934	34	RABMEMGLY	Oryctolagus cuniculus	9.28e+00	
c	24	132	9.3	43449	27	CELC18A11	Caenorhabditis elegans	9.28e+00
25	130	9.2	1599	8	BNPGALACR	B.napus mRNA for poly	1.41e+01	
c	26	131	9.2	2121	42	POABETA2	P.agilis beta-2 tubul	1.14e+01
27	130	9.2	6504	27	AAU02548	Aedes aegypti UGALS v	1.41e+01	
28	130	9.2	8780	31	MQSVITE	Aedes aegypti (clone	1.41e+01	
29	131	9.2	110908	26	SYCSLRD	Synechocystis sp. slr	1.14e+01	
30	128	9.0	1477	22	LEP1S1533	Leptospira borgpeters	2.13e+01	
c	31	127	8.9	684	47	CJDNASRY	C.jacchus SRY gene.	2.61e+01
32	127	8.9	830	26	TRNTETE	Transposon Tn10 (from	2.61e+01	
33	126	8.9	1337	79	XLTF3A36	Xenopus laevis gene e	3.19e+01	
c	34	127	8.9	2074	41	LEPOOXE	L.esculentum gene for	2.61e+01
c	35	127	8.9	2146	43	S40548	polyphenoloxidase [ly	2.61e+01
c	36	127	8.9	2214	40	EGOCNDHF	Erigeron hybridus chl	2.61e+01
c	37	127	8.9	2226	41	HAYCPNDHF	Haplocarpha scaposa c	2.61e+01
c	38	127	8.9	2226	39	ARZCPNDHF	Arctotis stochadifoli	2.61e+01
c	39	127	8.9	2772	18	AVINIFANFA	A.vinelandii nitrogen	2.61e+01
c	40	126	8.9	4032	53	HUM23DC1Z	Homo sapiens (subclon	3.19e+01
c	41	127	8.9	7194	50	HSENO3	H.sapiens ENO3 gene f	2.61e+01
c	42	127	8.9	10603	75	PAMFNL	Sendai virus (strain	2.61e+01
c	43	127	8.9	20235	25	SERERYAB	S.erythraea second an	2.61e+01
c	44	127	8.9	20444	25	SEERYABS	S.erythraea eryA gene	2.61e+01
c	45	126	8.9	40780	4	CEM03C11	Caenorhabditis elegans	3.19e+01

ALIGNMENTS

RESULT	1	HSU06654	752 bp	mRNA	PRI	30-JUL-1994
LOCUS		Human differentiation antigen melan-A protein mRNA, complete cds.				
DEFINITION		U06654				
ACCESSION		g517022				
NID		human.				
KEYWORDS		Homo sapiens				
SOURCE						
ORGANISM						



Qy 181 W:

Qy	121	YTYNTNATHGGNTGTGTATYTCYMGNMGMNAAAYGNTAYMNGNCNYTNATGGYAAAR	180
Db	234	agtcttcattgtgcaactcaatgtgccttaacaagaagatgcccaagaagggttgat	293
Qy	181	WSNYTNCAYCTGGNACNARTCYGNYTNACNMGMNGTGYCCNARGCGGNTTYGAY	240
Db	294	catcggcagacaaagtgtcttccaagagaaaaactgtgaacctgtgtcccaatgct	353
Qy	241	CAYMGNCAYWSNARGTWSNYTNCARGAARAATYCGARCCNTGTCNGCNAAACN	300
Db	354	ccacctgttatgaaactctctcagaacagtcaccaccacattatcacc	406
Qy	301	CNCNCNGTAYGARAARYTNWSNGNCARGAWSNCCNCCNCONATWSNCC	353
RESULT	3		
LOCUS	S77127	12857 bp	DNA
DEFINITION	manganese superoxide dismutase/MnSOD [human, embryoniclung fibroblast cell line W138, Genomic, 12857 nt].		
ACCESSION	S77127		
NTD	g998582		
KEYWORDS	human embryoniclung fibroblast cell line W138.		
SOURCE	Homo sapiens		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 12857)		
AUTHORS	Wan, X.S., Devalaraja, M.N. and St Clair, D.K.		26-SEP-1995
TITLE	Molecular structure and organization of the human manganese superoxide dismutase gene		
JOURNAL	DNA Cell Biol. 13 (11), 1127-1136		(1994)
MEDLINE	95217333		
REMARK	GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 164378] from the original journal article. This sequence comes from Fig. 3.		
COMMENT	Map location: 6.		
FEATURES	NCBI gi: 998582		
source	Location/Qualifiers		
	1..12857		
	/organism="Homo sapiens"		
	/note="human"		
mrna	join(857..881,1179..1363,5759..5876,8960..9140,11351..11498)		
	/note="Description: manganese superoxide dismutase/MnSOD"		
CDS	join(857..881,1179..1363,5759..5876,8960..9140,11351..11498)		
	/partial		
	/note="coding region is determined by the author; Description: manganese superoxide dismutase/MnSOD"		
	/codon_start=1		
	/translation="MLSRAVCGSGFGVSGAQAEQPPRALIRPCTSHQRADHAA		
	FQQAPRGREQPERHRRGEVGGVGGGMQLPR"		
BASE COUNT	3335 a 2627 c 3147 g 3748 t		
ORIGIN			
Query Match	12.1%;	Score 172;	DB 59; Length 12857;
Best Local Similarity	40.3%;	Pred. No. 9,32e-04;	
Matches	31; Conservative	20; Mismatches 26;	Indels 0; Gaps 0;
Db	12046	agcagdtgaactggggtgtctcactgcagtaagaacactcaacgcaa	12105
Qy	214	MGMNGTCCNARGCGGNTTYGAYCAYMNGAYWSNARGTWSNYTNCARGAAR	273
Db	12106	aactgcataaaggatggt	12122

QY	274	AAVTGTCARCONGTNGT 290	:  :				
RESULT	4						
LOCUS	ACU12386	41591 bp	DNA	circular INV 24-AUG-1995			
DEFINITION	Acanthamoeba castellanii mitochondrion, complete genome.						
ACCESSION	U12386						
NID	g562028						
KEYWORDS	23S-like ribosomal RNA; YMF46; YMF47; YMF48; Ile-tRNA; Ala-tRNA; Pro-tRNA; Asp-tRNA; initiator Met-tRNA; 16S-like ribosomal RNA; cytochrome oxidase subunit 1; cytochrome oxidase subunit 2; ribosomal protein 94; Phe-tRNA; X-tRNA; NADH dehydrogenase subunit 6; NADH dehydrogenase subunit 5; NADH dehydrogenase subunit 11; cytochrome oxidase subunit 3; NADH dehydrogenase subunit 4; NADH dehydrogenase subunit 2; ribosomal protein 52; His-tRNA; Leu-tRNA; H(+)-transporting ATPase subunit 9; Tyr-tRNA; cytochrome b; NADH dehydrogenase subunit 4L; ORF83; ORF115; H(+)-transporting ATPase subunit 1; NADH dehydrogenase subunit 1; Trp-tRNA; ribosomal protein 111; ribosomal protein 512; ribosomal protein 57; ribosomal protein 12; ribosomal protein 519; ribosomal protein 53; ribosomal protein 116; ribosomal protein 114; ribosomal protein 15; ribosomal protein 514; ribosomal protein 58; ribosomal protein 16; ribosomal protein 513; ribosomal protein 511; ORF5; ORF25; NADH dehydrogenase subunit 3; NADH dehydrogenase subunit 9; NADH dehydrogenase subunit 7; H(+)-transporting ATPase subunit 6; ORF349; Gln-tRNA; Lys-tRNA; Glu-tRNA.						
SOURCE	amoeba.						
ORGANISM	Mitochondrion Acanthamoeba castellanii Eukaryotae; mitochondrial eukaryotes; Acanthamoebidae; Acanthamoeba.						
REFERENCE	1	(bases 1 to 41591)					
AUTHORS	Burger, G., Plante, I., Loneragan, K.M. and Gray, M.W.						
TITLE	The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba castellanii. Complete sequence, gene content and genome organization						
JOURNAL	Unpublished						
REFERENCE	2	(bases 1 to 7352)					
AUTHORS	Loneragan, K.M. and Gray, M.W.						
TITLE	The ribosomal RNA gene region in Acanthamoeba castellanii mitochondrion DNA. A case of evolutionary transfer of introns between mitochondria and plastids?						
JOURNAL	J. Mol. Biol. 239 (4), 476-499 (1994)						
MEDLINE	94275838						
REFERENCE	3	(bases 5124 to 5757)					
AUTHORS	Loneragan, K.M. and Gray, M.W.						
TITLE	Editing of transfer RNAs in Acanthamoeba castellanii mitochondria						
JOURNAL	Science 259 (5096), 812-816 (1993)						
MEDLINE	93157849						
REFERENCE	4	(bases 1 to 7352)					
AUTHORS	Loneragan, K.M. and Gray, M.W.						
TITLE	Predicted editing of additional transfer RNAs in Acanthamoeba castellanii mitochondria						
JOURNAL	Nucleic Acids Res. 21 (18), 4402 (1993)						
MEDLINE	94021401						
REFERENCE	5	(bases 5124 to 5757)					
AUTHORS	Loneragan, K.M. and Gray, M.W.						
TITLE	Subunits 1 and 2 of cytochrome c oxidase are expressed from a continuous open reading frame in the mitochondrial DNA of Acanthamoeba castellanii						
JOURNAL	Unpublished						

## REFERENCE 6 (bases 1 to 41591)

AUTHORS Burger G.  
TITLE Direct Submission  
JOURNAL Submitted (16-JUL-1994) Gertraud Burger, University of Montreal, Biochemistry, 2900 Boul. Edouard-Montpetit, Montreal, Quebec, H3C 3J7, Canada

COMMENT NCBI gi: 562028

## FEATURES

source

Location/Qualifiers

1..41591

/mitochondrion

/strain="Neff (ATCC 30010)"

/organism="Acanthamoeba castellanii"

join(37..1806,2562..2581,3442..4074,4828..5123)

/gene="rnl"

/note="large subunit (LSU) ribosomal RNA"

/citation=[1]

/citation=[2]

/product="23S-like ribosomal RNA"

37..1806

/gene="rnl"

/note="large subunit (LSU) ribosomal RNA"

/number=1

/citation=[1]

/citation=[2]

/product="23S-like ribosomal RNA"

1807..2561

/note="group I intron, AcLSU.ml"

/number=1

/citation=[1]

/citation=[2]

2012..2440

/gene="ymf46"

/note="rnl intron 1 (AcLSU.ml) orf, group I orf; NCBI gi:

562029"

/codon\_start=1

/citation=[1]

/citation=[2]

/transl\_table=4

/product="YMF46"

/db\_xref="PID:g562029"

/translation="MKIDQWIVGVGEGCFYICINKSVDSKLGQYVPEERVQHK

RDIKLYIAKDOFFGHGSVVCNKSNGSEIYEYRVRKFETLHDVILPFFESNGLTSKX

NFLAFRDVILMKREHLTESGLSKIIDIKSMNRSIH"

2562..2581

/gene="rnl"

/number=2

/citation=[1]

/citation=[2]

2582..3441

/note="group I intron, AcLSU.m2"

/number=2

/citation=[1]

/citation=[2]

2849..3355

/gene="ymf47"

/note="rnl intron 2 (AcLSU.m2) orf, group I orf; NCBI gi:

562030"

/codon\_start=1

/number=2

/citation=[1]

/citation=[2]

/transl\_table=4

/product="YMF47"

/db\_xref="PID:g562030"

/translation="MNLDLQKQKALIDPFWISGFVDGEGCFICISFNKERLTGLIEV  
RPSFSISQTRDKEGIALRCIADFLNFFDCGFIFRSKRDNTWKYECRDLSDIRSKVLEPH  
FEFETLRTKLRDFELFKDVVNSVASKQHINEVGLKRIIDISYQINMGKKLTKDELL  
SKINLNKL"

exon

/gene="rnl"

/number=3

/notes="group I intron, AcLSU.m3"

/number=3

/citation=[1]

/citation=[2]

4207..4701

/gene="ymf48"

/note="rnl intron 3 (AcLSU.m3) orf, group I orf; NCBI gi:

562031"

/codon\_start=1

/number=3

/citation=[1]

/citation=[2]

/transl\_table=4

/product="YMF48"

/db\_xref="PID:g562031"

/translation="MOKNQFKLNQQLAYLAGFEADGCFLQIIPGLQYRYKHTIR

ISIVFYQKDKHWYFLQIAKLNILIGSIRFNDGMLEYITGLSVNKFLEMLFPVLL

KKNLAVLIPRIIKGLDVRNENGFLVCKLVDEVADHTYSKRRKNTSLTVKNSLLLPV

ETEE"

4828..5123

/gene="rnl"

/number=4

/citation=[1]

/citation=[2]

5157..5227

/gene="trnI2"

/note="Gene originally designated trnM1 [2], [3]; now

identified as a homolog of the E. coli gene encoding an

Ile-tRNA with anticodon LAU, where L is lysidine, a

modified C"

/anticodon=(pos:unspecified,aa:Ile)

/citation=[1]

/citation=[2]

/citation=[3]

/product="Ile-tRNA"

5258..5328

/gene="trnA"

/anticodon=(pos:unspecified,aa:Ala)

/citation=[1]

/citation=[2]

/citation=[3]

/product="Ala-tRNA"

5336..5408

/gene="trnP"

/anticodon=(pos:unspecified,aa:Pro)

/citation=[1]

/citation=[2]

/citation=[3]

/product="Pro-tRNA"

5412..5484

/gene="trnD"

/anticodon=(pos:unspecified,aa:Asp)

/citation=[1]

/citation=[2]

/citation=[3]

/product="Asp-tRNA"



[illegible]

RESULT	7	LEPREPEL	1467 bp	DNA	BCT	01-OCT-1991
LOCUS		Leptospiira hardjo				
DEFINITION		repetitive element.				
ACCESSION		M57713				
KEYWORDS		repeat region.				
ORGANISM		Leptospiira hardjo				

apicomplexes), Leptospiraceae.  
1 (bases 1 to 1467)  
Woodward, M.J. and Sullivan, G.J.  
Nucleotide sequence of a repetitive element isolated from  
*Leptospira interrogans* serovar hardjo type hardjo-bovis  
J. Gen. Microbiol. 137, 1101-1109 (1991)  
JOURNAL

MEDLINE	91324852
COMMENT	NCBI gi: 149621
FEATURES	Location/Qualifiers
	1..1467
	source

BASE COUNT	526 a	240 c	338 g	363 t

Query Match 9.7%; Score 138; DB 22; Length 1467;  
Best Local Similarity 33.7%; Pred. No. 2.58e+00;  
Matches 35; Conservative 26; Mismatches 42; Indels 1; Gaps 1;

Db 589 ggttgagtcattacaaatggttgaacaatcttcactttaacaacgagatccttcaagaga 648

Qy 174 GGAYARWSYTNCA YGTNGGNACNCARTGYGCNYTNACNMGNMGNTGYCCN-CARGARG 232

Db 649 cgttaacgattattatagtccgactcgaagattccaagaacagaat 692

233 GNTTYGAYCMGNCAYMSNAARGTWSNYTNCARGAABAAY 276

RESULT	8	HSGLNRS09	267 bp	DNA	PRI -	11-MAR-1993
LOCUS						
DEFINITION						

ACCESSION X72404

NID  
g287775

**KEYWORDS** glutamyl-tRNA synthetase.

SOURCE	human.
ORGANISM	human.

ORGANISM

Homo sapiens

Eukaryotes; mitochondrial eukaryotes; Metazoa/Eumycota group;

Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;

Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;

Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;

Catarrhini; Hominoidea; Homo

REFERENCE  
AUTHORS  
TITLE

1 (bases 1 to 267)  
Kaiser, E., Eberhard, D. and Knippers, R.  
Exons encoding the highly conserved part of human glutamyl-tRNA synthetase

J. Mol. Evol.	34 (1), 45-53 (1992)
JOURNAL MEDLINE REFERENCE	92211721
AUTHORS	2 (bases 1 to 267) Kunze, N., Bittler, E., Fett, R., Schray, B., Hameister, H., Wiedorn, K. H. and Knippers, R.
TITLE	The human QARS locus; assignment of the human gene for glutamyl-tRNA synthetase to chromosome 1q32-42
JOURNAL	Hum. Genet. 85 (5), 527-530 (1990)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

3 (bases 1 to 267)  
Kaiser, E.  
Direct Submission  
Submitted (26-FEB-1993) to the EMBL/GenBank/DBJ databases. E.  
Kaiser, AG Prof. Dr. R. Knippers, Institut f. Mol. Genetik, Univ.  
Konstanz, Fakultät fuer Biologie, Universitätsstr. 10 D-7750

COMMENT	NCBI gi: 287775
FEATURES	Location/Qualifiers
source	1..267

```

/chr_name="1"
/chromosome="1"
/map="1q32-42"
<1..175
/number=8
176..260
/number=9
intron
exon
intron
/number=9

```

Query Match 9.6%; Score 137; DB 50; Length 267;  
Best Local Similarity 37.3%; Pred. No. 3.21e+00;  
Matches 25; Conservative 17; Mismatches 25; Indels 0; Gaps 0;

Db 14 tctgattggctcctagtattgcatactacctaattagtataaactagctcttagacaaat 73

[illegible]

QY 122 TNYTNATHGCGTGTGGTAYTGYMGNMGMGNAAYGCGNTAYMCGNGCNYTNATGGAYARW 181

74 80

DD : : ||

Qy 182 SNYTNCA 18

RESULT	9	DATE	TIME
LOCUS	CMDD21	1204	12
			1005

LOCUS CIBRZ1 1384 bp RNA INV 13-JUL-1993  
DEFINITION Chironomus tentans Balbiani ring mRNA BR 2.1 3'-end.

ACCESSION X07701

NID q7053

**KEYWORDS** BR2.1 gene.

SOURCE Chironomus tentans.

ORGANISM Chironomus tentans

Eukaryota; Animalia; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Holometabola; Diptera; Nematoxya; Culicidae;

Chironomidae.  
Fleurygata, Neoptera, Holometabola, Diptera, Nematocera, Culicorpha.

REFERENCE 1 (bases 1 to 1384)

**AUTHORS** Hoog, C., Daneholt, B. and Wieslander, L.

**TITLE** Terminal repeats in long repeat arrays are likely to reflect the

early evolution of Balbiani ring genes

J. MOL. BIOL. 200 (4), 653-664 (1986)  
JOURNAL MEDLINE 88316932

COMMENT NCBI qi: 7053

**FEATURES**

**Location/Qualifiers**



```
/dev stage="embryonic (3-15 h)"
/clone lib="embr. cDNA"
/clone="C10.1.1, C10.2.1, C2.3.1, E4.1, E7.1"
158..217
sig_peptide

CDS
/gene="Notch"
/notes="transmembrane protein"
158..7471
/gene="Notch"
/notes="NCBI gi: 433867"
/codon start=1
/product="transmembrane protein-precursor"
/db xref="PIR:g433867"
/translation="MNRFLVKLITLTAASLATVAQRCQSEYQNGGICEYKPSGEAS
CRPADFGCAQCFPNPQNPSPCRNGVGRPMQMGVGVCDVGLGFSRLCLTPVN
HACMNSPQRNGGTSLLITDFTFCRQPCGMSKTCQLADPCASNPANGCQGSASFES
YICTPPNPHGQTCQDVNCAVASPFCRNGGTCNEVGSYLCRPPETYPGPQRLY
QPCLPSPCRSGGTCVSTDTHTGSLPGFTGOTCEHNVDDCTQHACENGRCIDGIN
TYNCHCKWGTGQCTEDVDECELSNPAQNGGTCNHTTGGFHCVCVNGTGDGSEN
IDDCASAGCSHGTCDHRVASFFCEPHGRTGLLCHLDACISNPQKSGNDYNPVS
GKAICTPPGVTSAGNODIDECISLANPCEHGRCLNTKGSFOCKCLAQCYGPRCM
DVNEKSNPQNDATCLDQIGGFHCICMFGYGVFCQINSDDCASQPCINGKCIDKN
SFHCPCPGFSGSCQVDVDECASTPCRNKAGCTDGPNTYCECTPFGSGHCELDIN
ECASSPCHVGRGVASTCDRPGVYTRGLCTNINELSQPCRNNGTCQDRENAVYI
CTCPKGTGVCNENIDDCRKPCKDKINGECYCEPFGYSGSKNINIDDCAL
NPNCHGCTGIDVNSFTCLCPDFRDTCLDSHNECSNPCIHGSCLDQINSYRCVE
AGWNGRNDININELSNPVNGGTCMDTSGYLCRAGFSGNQMINECASNP
LNQSGCIDVAGKCNCLPYYGVCEVNLAPSPRCNKGVCRESDFQFSNCP
AGAGQTCVDINCEVNPCTNGGVENLRGFCRNPFTGALCENDIDDCENPC
SNGVGRVNGFVYCLAGRCVCAEDIDECVSPCRNGGCTDCVNSYTCSPAG
FSGINCEINTDCTESSCNGCTGDCISFSCYCLPFTGNYCOHDVNECDRQCM
GSGQDGYTKTCFPHGTGLNCSLVNDCSSPCRNQSGCQMGASFTQCAAGMT
GTYDVPVSVECAARQGVSAVLCHAGQCVAGNTHLCRQAGYTGSCYQEQVDE
CQNPQNGCATDYLGGYSCVCPGYHGMNCSKEINELSQPCNGGTCIDLNTYK
CSPRGTQVHCEIDDDCSPVDP LFGPCFNGRCVDRVGGVGVCPGFGVERC
EGDVNCLSDPCDP SGYNVOLINDFRCEGTGTGKRCETVFNCKDTPCKNGTC
AVASNTKHCYCKCQPCVSGCYSDSCGSLRCRNGATCVSGHLSPRCLAPGFCG
HECQTRMDSPLVNPVNGGTCQPIDSAFYRCSPANFNLLCHILIDYFSFGQGRD
IAPVVEVEIRCEIAAGCEGLYDGCQRLEGCNP LYDQYCRDHYADGCHDCGNNAEC
WRYDNGKQDEQACATGCLYDGCQRLEGCNP LYDQYCRDHYADGCHDCGNNAEC
EMDGLDCAADVPQKLVGSLVHVIIDPDELNRSSSF LNELSLHTNVTFRDANG
EALIFPYYSGEHLSKHKRSMDTPQGLAQRRSLTSFLAPTRRELDHMEVGSIV
YLEIDNRQCFQSDCEFSATDVAAFLGALASSGNLNVPI TEAVTSEGGPPTGEMV
PMFLVLLALAVLAAVGVVSVRRKREHQGLWPFEGFKVNEPKKRRREPVEDSVGL
KP LKNSDS LMDQGLSEWADDYTKRFRFEGQSILEMSQGLDHRQMTQQLDAADJRL
NSMAPTPPOQIENDCMVNVGPDGFTPLMTASCSGGGLENEGEAEEDPSADVTID
FIYHGANLHQGTGTGETALHAAARYAARDAKRLLESCADANQDMNGRTP LHAVA
ADAGQVQLILNRNDLDRHIDGTTPLILATRLAVEGMVEELINCHADPNVDDSG
KSA LHAWAANNVDAVAVLLKNGKNDLONKEETPLFLAAREGSYETAKVLLHLAN
RDIAHDJQLPRDIAHERMHHDIYVLLLEYNLVRSPP LPLSPPLCCPNTYLKTFSPG
NNNNTAKTRPFGKPGKGGKSGKD ITRKKKSGDKNGKMEGVLSPLSLESHPG
YLSNVSPSPMPTSPFOQSPPI SIALQGLADSHMGALQGLKRPDPSAPRLSLHPVAN
NVGGAQAGACWILQRVQOQQOQQOQQOQAGMLPMTLSA NMPQVNGYPTMSSHLGA
P SHMIAHQMAPMQHNT SHHF LGLDLSGLDIQSSSGHAP IQTILPDQSBNAPP LLSST
QFLTPPSQHSYSNPMNDTNRHQVDPDHPFLTP SAGSPDQWSSSPSHSLNLSDEGIS
PPTSMQWNIPEAFK"
mat_peptide
218..7468
/gene="Notch"
/product="transmembrane protein"
BASE COUNT 1726 a 2046 c 2104 g 1595 t
ORIGIN
```

Query Match 9.6%; Score 137; DB 77; Length 7471;  
Beat Local Similarity 31.3%; Pred. No. 3.21e+00;

```
Matches 61; Conservative 36; Mismatches 97; Indels 1; Gaps 1;
Db 2646 cttgtctccagtcctgtaaaaacgqcggtgtgtcgtgaatctgaggtattccaga 2705
::||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Cp 334 SYTGTCNGMNSNARYTTCRTTANGCNGNGGCGRTTNGNACNACNGGTYCRCA 275
::||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2706 gcttcctctgaactgtccgctggatggcaa-ggtcaaacgtgtgaggtgacatta 2764
::||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Cp 274 TTTTTCYTCGNARNWSNACYTTNWSRTCNCKTCTCRAANCCTCTCTGNGRCAN 215
::||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2765 gagtgtgtgaggaacccgtgactaataggagagtggtgaaacacgtgcgggtgatt 2824
::||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Cp 214 KNGTNARNGRCAYTNGTNGCNCACRTGNARNSYTRTCCATNARNGCCKRTANC 155
::||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2825 cagtgcgcgtgcaat 2839
::||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Cp 154 TNCNCKNCKRCART 140
::||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
RESULT 12
LOCUS CELF47F2 24537 bp DNA INV 25-NOV-1995
DEFINITION Caenorhabditis elegans cosmid F47F2.
ACCESSION U04943
NID g1072202
KEYWORDS
SOURCE Caenorhabditis elegans strain-Bristol N2.
ORGANISM Caenorhabditis elegans
Eukaryotes; mitochondrial eukaryotes; Metazoa; Nematoda;
Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea;
Rhabditidae; Caenorhabditis.
REFERENCE
1 (bases 1 to 24537)
Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M.,
Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J.,
Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favell, A.,
Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M.,
Johnston, L., Jones, M., Kersey, J., Kirsten, J., Laister, N.,
Latreille, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B.,
O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Roopra, A.,
Saunders, D., Showkneen, R., Smaildon, N., Smith, A., Sonhammer, E.,
Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M.,
Vaughan, K., Waterston, R., Watson, A., Weinstock, L.,
Wilkinson-Sproat, J. and Wohldman, P.
2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans
```

```
JOURNAL Nature 368 (6466), 32-38 (1994)
MEDLINE 94150718
REFERENCE 2 (bases 1 to 24537)
Bentley, D.
TITLE The sequence of C. elegans cosmid F47F2
JOURNAL Unpublished (1995)
REFERENCE 3 (bases 1 to 24537)
Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-1995) Robert Waterston
COMMENT Submitted by:
Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England
e-mail: rwenematode.wustl.edu and jee@sanger.ac.uk
NEIGHBORING COSMID INFORMATION:
5' cosmid is 2C64, 200 bp overlap; 3' cosmid is T22B2. Actual start
of this cosmid is at base position 29748 of CELF2C64.
```





RESULT 15  
LOCUS OATCVG6 708 bp RNA MAM  
DEFINITION O.aries rearranged T-cell receptor gamma. 15-FEB-1993

Qy 280 GARCCNGTNGTNC 293

Search completed: Tue Aug 27 18:41:38 1996  
Job time : 276 secs.

\*\*\*\*\*

(TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Aug 27 08:08:49 1996; MasPar time 3.69 Seconds  
282.669 Million cell updates/sec

Tabular output not generated.

Title: >US-08-231-565A-2

Description: (1-118) from US08231565A.pep

Perfect Score: 889

Sequence: 1 MPREDAHFTYCPKKGCHS.....NAPPAYEKLAEQSPPPYSP 118

Scoring table: PAM 150  
Gap 11

Searched: 76839 seqs, 8833555 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq23

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15

Statistics: Mean 29.335; Variance 113.395; scale 0.259

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	889	100.0	118	15	R84212	MART-1 melanoma antiq 1.71e-83
2	889	100.0	118	11	R63158	Tumour rejection anti 1.71e-83
3	90	10.1	380	1	R05433	cPA-P2 Hybrid plasmin 3.71e+00
4	83	9.3	200	15	R82900	Mouse B7-1 (IgV-like 1.37e+01
5	83	9.3	212	15	R82902	Mouse B7-1 IgV-like i 1.37e+01
6	83	9.3	306	15	R82893	Mouse B7-1 alternati 1.37e+01
7	83	9.3	306	12	R67990	Murine B lymphocyte a 1.37e+01
8	82	9.2	897	4	R20982	Sequence of beta-chai 1.65e+01
9	81	9.1	455	5	R28757	Hepatocyte nuclear fa 1.98e+01
10	81	9.1	855	10	R51252	FIV UK2 envelope prot 1.98e+01
11	81	9.1	856	10	R51248	FIV petaluma envelope 1.98e+01
12	81	9.1	856	10	R51249	FIV PET-F14 envelope 1.98e+01

13	81	9.1	856	10	R51247	FIV envelope protein	1.98e+01
14	79	8.9	234	3	R13844	Fusaric acid resistan	2.84e+01
15	79	8.9	854	10	R51250	FIV PPR envelope prot	2.84e+01
16	79	8.9	856	10	R51254	FIV DUTCH19K1 envelop	2.84e+01
17	79	8.9	940	2	R07070	Fibronectin-binding p	2.84e+01
18	76	8.5	524	2	R05077	Sali restriction prod	4.86e+01
19	76	8.5	524	1	P90113	Rabies G protein	4.86e+01
20	76	8.5	854	8	R44957	Feline Immunodeficien	4.86e+01
21	76	8.5	1822	10	R55273	Beta subunit of integ	4.86e+01
22	75	8.4	117	2	R07997	Human CD2 cytoplasmic	5.80e+01
23	75	8.4	332	13	R74221	Epitope on the primar	5.80e+01
24	75	8.4	345	15	R77026	Canine herpesvirus gd	5.80e+01
25	75	8.4	351	4	R20803	Human CD2 antigen.	5.80e+01
26	75	8.4	351	2	R11921	Human T11 sugar prote	5.80e+01
27	75	8.4	353	2	R06365	Soluble two domain hu	5.80e+01
28	75	8.4	360	1	P81178	Sequence of human T11	5.80e+01
29	75	8.4	855	10	R58585	FIV UK8 envelope prot	5.80e+01
30	75	8.4	855	10	R51255	FIV SWISS2 envelope	5.80e+01
31	75	8.4	870	9	R51246	Composite sequence of	5.80e+01
32	74	8.3	132	7	R36393	Penetrin.	6.91e+01
33	74	8.3	1026	11	R04924	Immunoprotein TANG12.	6.91e+01
34	74	8.3	1026	11	R58906	Human protocaderhin-4	6.91e+01
35	74	8.3	1203	11	R58911	Product of alternativ	6.91e+01
36	73	8.2	10	15	R84778	MART-1 melanoma antiq	8.23e+01
37	73	8.2	238	2	P70375	Human interleukin-2 r	8.23e+01
38	73	8.2	272	3	P50001	Sequence of interleuk	8.23e+01
39	73	8.2	272	14	R79389	Human interleukin-2 r	8.23e+01
40	73	8.2	272	3	P61052	Sequence encoding int	8.23e+01
41	73	8.2	896	4	R21580	Plasma membrane proto	8.23e+01
42	72	8.1	10	15	R84777	MART-1 melanoma antiq	9.80e+01
43	72	8.1	479	1	P81013	Complete sequence of	9.80e+01
44	72	8.1	479	3	P50034	Sequence encoded by t	9.80e+01
45	72	8.1	1091	2	P98500	Partial sequence enco	9.80e+01

## ALIGNMENTS

RESULT	1
ID	R84212 standard; Protein; 118 AA.
AC	R84212;
DT	20-APR-1996 (first entry)
DE	MART-1 melanoma antigen.
KW	MART-1; melanoma antigen recognised by T-cell; melanoma;
KW	metastatic melanoma; tumour-associated antigen; immunogen;
KW	diagnosis; prognosis; prophylaxis; therapy; vaccine.
OS	Mammalian.
EH	Key
FT	Region
FT	/note= "hydrophobic region"
PN	WO9529193-A2.
PD	02-NOV-1995.
PF	21-APR-1995; U05063.
PR	22-APR-1994; US-231565.
PR	05-APR-1995; US-417174.
PA	(USSH ) US SEC DEPT HEALTH.
PI	Kawakami Y, Rosenberg SA;
DR	WPI; 95-382963/49.
DR	N-PSDB; T02714.
PT	DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT	vectors, host cells and antibodies, used to detect, treat and
PT	immunise animal against melanoma.
PS	Claim 11; Page 117; 184pp; English.
CC	The melanoma antigen (MART-1) is produced by recombinant DNA
CC	methods, i.e. preferably using a baculovirus vector for expression

CC in insect cell cultures. MART-1 protein is a source of immunogenic peptides (see R84196 for peptide M9-2) which are optionally modified (see R84783-R84800) and used in medicaments for the treatment or prevention (by immunization) of melanoma. Antibodies against MART-1 and its immunogenic peptides may be used in the detection and isolation of MART-1 from a sample, the detection of which is indicative of a disease state (melanoma or metastatic melanoma).  
SQ Sequence 118 AA;

Query Match 100.0%; Score 889; DB 15; Length 118;  
Best Local Similarity 100.0%; Pred. No. 1.71e-83;  
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 1 mpreahfgygpkqghgheyttaeagilgtvlllgwycwrrrrngyralmdk 60  
Qy 1 MPRDAHFIYGYPKKGHGYTTAEAGIGILTVILGVLLIGCWYCRRRNGYRALMDK 60  
Db 61 slhvtgcaltrrcpqegfhdhdsksvlqekncepvvpnapppayeklaeqspppy 118  
Qy 61 SLHVGTCALTRCPQEGFHDHDSKSVLQEKNCPEVPVNPAPPAYEKLAEQSPPPY 118

# RESULT 2

ID R63158 standard; Protein; 118 AA.  
AC R63158;  
DT 26-MAY-1995 (first entry)  
DE Tumour rejection antigen precursor.  
KW Tumour rejection antigen; precursor; HLA-A2 molecule; tyrosinase; isolation; melanoma; cell line; LB-39-MEL; diagnosis; vaccine; therapy.  
OS Homo sapiens.  
PN W09421126-A.  
PF 29-SEP-1994.  
PR 09-MAR-1994; U02487.  
PR 18-MAR-1993; US-032978.  
PI (LUDM-) LUDMIG INST CANCER RES.  
PI Boon-Fallieur T, Brichard V, De Plaen E, Traversari C;  
PI Van Pel A, Wolfel T;  
DR WPI; 94-316544/39.  
DR N-PSDB; Q76370.  
PT Nucleic acid coding for a tumour rejection antigen precursor - is used for developing prods. for diagnosis or treatment of expression related disorders, partic. melanoma  
PS Claim 5; Page 14; 26pp; English.  
CC This sequence represents the tumour rejection antigen precursor which is processed to a tumour rejection antigen presented by HLA-A2 molecules. The tumour rejection antigen is not related to tyrosinase. The cDNA encoding this sequence was isolated from the melanoma cell line, LB-39-MEL. The tumour rejection antigen may be used for diagnosis or in vaccines or for therapy of disorders characterised by the expression of the tumour rejection antigen precursor, particularly melanoma.  
SQ Sequence 118 AA;

Query Match 100.0%; Score 889; DB 11; Length 118;  
Best Local Similarity 100.0%; Pred. No. 1.71e-83;  
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mpreahfgygpkqghgheyttaeagilgtvlllgwycwrrrrngyralmdk 60  
Qy 1 MPRDAHFIYGYPKKGHGYTTAEAGIGILTVILGVLLIGCWYCRRRNGYRALMDK 60  
Db 61 slhvtgcaltrrcpqegfhdhdsksvlqekncepvvpnapppayeklaeqspppy 118  
Qy 61 SLHVGTCALTRCPQEGFHDHDSKSVLQEKNCPEVPVNPAPPAYEKLAEQSPPPY 118

RESULT 3  
ID R05433 standard; protein; 380 AA.  
AC R05433;  
DT 30-JUL-1990 (first entry)  
DE cPA-P2 Hybrid plasminogen activator.  
KW Plasminogen activator; fibrin; urokinase; thromboembolic disease; ds.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT Domain 1..23  
FT /label=Secretory leader.  
FT Domain 24..102  
FT /label=Plasminogen Kringle 1  
FT Domain 103..116  
FT /label=Urokinase linker  
FT Domain 117..380  
FT /label=ScuPA protease domain.  
PN W09001332-A.  
PD 22-FEB-1990.  
PF 10-AUG-1988; 02771.  
PR 10-AUG-1988; W0-U02771.  
PI (CETU) Cetus Corp.  
PI Halluin AP;  
DR WPI; 90-083374/11.  
DR N-PSDB; Q02301.

PT Compn. contg. plasminogen activator conjugated to heparin component - used for treatment of thromboembolic disease, with longer half life and improved targeting.  
PS Disclosure; p; English.  
CC Gene encodes hybrid plasminogen activator (PA) comprising Kringle 1, an urokinase linker, and an urokinase protease domain wherein glycine residue at position 158 is replaced with a lysine.  
CC The compound is used to treat thromboembolic disease esp. with myocardial infarction, has a longer half-life than free PA and targets the heparin site of thrombus or embolism reducing the risk of reocclusion.  
SQ Sequence 380 AA;

Query Match 10.1%; Score 90; DB 1; Length 380;  
Best Local Similarity 30.9%; Pred. No. 3.71e+00;  
Matches 21; Conservative 15; Mismatches 24; Indels 8; Gaps 7;

Db 7 lallllllllpgcwascktdgknygtmstkn-gitcqwststphrfsphps 65  
Qy 33 LTVILGVLLIGCW-Y-CRRNG--YRALMDKSLHVGTCAL-TRCPQEG-FDH-RDSK 85  
Db 66 egleenyc 73  
Qy 86 VSIQEKNC 93

# RESULT 4

ID R82900 standard; Protein; 200 AA.  
AC R82900;  
DT 07-MAY-1996 (first entry)  
DE Mouse B7-1 (IgV-like domain deleted).  
KW T-cell costimulatory molecule; B7-1; T-lymphocyte; CD28; CTLA4; receptor; immunoglobulin.  
OS Mus musculus.  
FH Key Location/Qualifiers  
FT Peptide 1..37  
FT /label= Sig\_peptide  
FT Domain 169..200

FT /label= Cytoplasmic domain  
 FT /note= "cytoplasmic domain is encoded by exon 5  
 FT of the B7-1 gene"  
 PN W09523859-A2.  
 PD 08-SEP-1995.  
 PF 02-MAR-1995; U02576.  
 PR 02-MAR-1994; US-205697.  
 PA (BGHM ) BRIGHAM & WOMENS HOSPITAL.  
 PA (DAND ) DANA FARBEN CANCER INST.  
 PI Borriello F, Freeman GJ, Nadler LM, Sharpe AH;  
 DR WPI; 95-320574/41.  
 DR N-PSDB; T01047.  
 PT Novel T cell co-stimulatory molecules - corresponding to naturally  
 PT occurring alternatively spliced forms of T cells co-stimulatory  
 PT molecules or variants  
 PS Disclosure; Page 55-56; 111pp; English.  
 CC A naturally occurring form of mouse T-cell costimulatory molecule  
 CC B7-1 (R82900) has the signal peptide directly linked to the  
 CC IgG-like domain, i.e. the IgV-like domain is deleted. It is  
 CC encoded by exons 1, 3, 4 and 5 (see T01047) of the B7-1 gene. An  
 CC alternatively spliced form of IgV-deleted B7-1 (R82901) is encoded  
 CC by exons 1, 3, 4 and 6. T-cell costimulatory molecules can be  
 CC produced in which the IgV-like domain is deleted.  
 SQ Sequence 200 AA;

Query Match 9.3%; Score 83; DB 15; Length 200;  
 Best Local Similarity 27.6%; Pred. No. 1.37e+01;  
 Matches 8; Conservative 14; Mismatches 6; Indels 1; Gaps 1;

Db 146 gagfgavttvvvviiko-fckhrscfr 173  
 Qy 27 AAGIGILTIVILGVLLIGCWYRRNGYR 55

RESULT 5  
 ID R82902 standard; Protein; 212 AA.  
 AC R82902;  
 DT 07-MAY-1996 (first entry)  
 DE Mouse B7-1 IgV-like isoform.  
 KW T-cell costimulatory molecule; B7-1; T-lymphocyte; CD28; CTLA4;  
 KW receptor; immunoglobulin; interleukin-2.  
 OS Mus musculus.  
 FH Key Location/Qualifiers  
 FT Peptide 1..37  
 FT /label= Sig\_peptide  
 FT Domain 181..212  
 FT /label= Cytoplasmic domain  
 FT /note= "cytoplasmic domain is encoded by exon 5  
 FT of the B7-1 gene"  
 PN W09523859-A2.  
 PD 08-SEP-1995.  
 PF 02-MAR-1995; U02576.  
 PR 02-MAR-1994; US-205697.  
 PA (BGHM ) BRIGHAM & WOMENS HOSPITAL.  
 PA (DAND ) DANA FARBEN CANCER INST.  
 PI Borriello F, Freeman GJ, Nadler LM, Sharpe AH;  
 DR WPI; 95-320574/41.  
 DR N-PSDB; T01049.  
 PT Novel T cell co-stimulatory molecules - corresponding to naturally  
 PT occurring alternatively spliced forms of T cells co-stimulatory  
 PT molecules or variants  
 PS Disclosure; Page 91-92; 111pp; English.  
 CC A naturally occurring form of mouse T-cell costimulatory molecule  
 CC B7-1 (R82902) has the IgV-like domain directly linked to the

CC transmembrane domain, i.e. the IgC-like domain is deleted. It is  
 CC encoded by exons 1, 2, 4 and 5 of the B7-1 (T01049) gene.  
 CC This IgV-like isoform of B701 was expressed in CHO cells. It  
 CC triggered a costimulatory signal in T-cells, causing stimulation  
 CC of interleukin-2 prodn.  
 SQ Sequence 212 AA;

Query Match 9.3%; Score 83; DB 15; Length 212;  
 Best Local Similarity 27.6%; Pred. No. 1.37e+01;  
 Matches 8; Conservative 14; Mismatches 6; Indels 1; Gaps 1;

Db 158 gagfgavttvvvviiko-fckhrscfr 185  
 Qy 27 AAGIGILTIVILGVLLIGCWYRRNGYR 55

RESULT 6  
 ID R82893 standard; Protein; 306 AA.  
 AC R82893;  
 DT 07-MAY-1996 (first entry)  
 DE Mouse B7-1 alternatively spliced form.  
 KW T-cell costimulatory molecule; B7-1; T-lymphocyte; CD28; CTLA4;  
 KW receptor; immunoglobulin.  
 OS Mus musculus.  
 FH Key Location/Qualifiers  
 FT Peptide 1..37  
 FT /label= Sig\_peptide  
 PN W09523859-A2.  
 PD 08-SEP-1995.  
 PF 02-MAR-1995; U02576.  
 PR 02-MAR-1994; US-205697.  
 PA (BGHM ) BRIGHAM & WOMENS HOSPITAL.  
 PA (DAND ) DANA FARBEN CANCER INST.  
 PI Borriello F, Freeman GJ, Nadler LM, Sharpe AH;  
 DR WPI; 95-320574/41.  
 DR N-PSDB; T01040.  
 PT Novel T cell co-stimulatory molecules - corresponding to naturally  
 PT occurring alternatively spliced forms of T cells co-stimulatory  
 PT molecules or variants  
 PS Disclosure; Page 49-51; 111pp; English.  
 CC An alternatively spliced form of murine T-cell costimulatory  
 CC molecule B7-1 (R82893) is encoded by exons 1-6 (T01040) of the B1-7  
 CC gene. It includes cytoplasmic domains (R82887 and R82888) encoded  
 CC by exons 5 (T01035) and 6 (T01034). Novel cytoplasmic forms of  
 CC T-cell costimulatory molecules have been produced that bind to  
 CC CD28 or CTLA4 and trigger a costimulatory signal in T-cells.  
 SQ Sequence 306 AA;

Query Match 9.3%; Score 83; DB 15; Length 306;  
 Best Local Similarity 27.6%; Pred. No. 1.37e+01;  
 Matches 8; Conservative 14; Mismatches 6; Indels 1; Gaps 1;

Db 252 gagfgavttvvvviiko-fckhrscfr. 279  
 Qy 27 AAGIGILTIVILGVLLIGCWYRRNGYR 55

RESULT 7  
 ID R67990 standard; Protein; 306 AA.  
 AC R67990;  
 DT 21-AUG-1995 (first entry)  
 DE Murine B lymphocyte antigen B7 (mB7).  
 KW B lymphocyte activation antigen; B7-1, Ig superfamily; CD28;  
 KW transmembrane protein.



CC region. The zinc finger domain is flanked by regions with no known  
 CC similarity, but there is a large hydrophobic region in the C  
 CC terminal half (133-373) which has definite similarity to the ligand  
 CC binding domain of other receptors (20-37% identity), with H-2RIIBP  
 CC being most similar at 37.3% identity. The protein also has a  
 CC proline rich region (23%) at the C terminus (400-477) which could  
 CC be an activator domain, and three serine/threonine rich regions  
 CC (30-38%) scattered through the molecule which may be  
 CC phosphorylation sites. It is not known if HNF-4 is modified,  
 CC but some post translational modification is suggested by a  
 CC molecular weight of 54kD by SDS-PAGE, but 50.6kD from the predicted  
 CC amino acid sequence. The protein itself, the gene encoding it,  
 CC Abs, and antidiabetic Abs may be used to develop diagnostic and  
 CC therapeutic agents to detect, inhibit or enhance binding to HNF-4  
 CC They can be used to study, diagnose, prevent and treat diseases  
 CC such as coronary heart disease, hyperlipidaemia, liver disease and  
 CC arteriosclerosis. They may also be used in the treatment of  
 CC obesity.  
 CC Sequence 455 AA;

Query Match 9.1%; Score 81; DB 5; Length 455;  
 Best Local Similarity 30.4%; Pred. No. 1.98e+01;  
 Matches 14; Conservative 11; Mismatches 20; Indels 1; Gaps 1;

Db 103 crycrlkckfragnkkaevqnerdrietrssyedislpainallq 148  
 :|||| : ||| : : : ||| : : : ||  
 Qy 45 CMYCRRRNGYALMDK-SLAVGTQCALTRRCQEGFDRHDSKVSILQ 89

RESULT 10  
 ID R51252 standard; peptide; 855 AA.  
 AC R51252;  
 DT 08-OCT-1994 (first entry)  
 DE FIV UK2 envelope protein sequence.  
 KW Feline immunodeficiency virus; FIV; vaccine; diagnostic; AIDS;  
 KW T-lymphotropic lentivirus; FIV UK8; FIV Petaluma; envelope protein.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..178  
 FT /note= "Hydrophobic leader peptide"  
 FT Region 51..66  
 FT /note= "V1 region"  
 FT Region 95..173  
 FT /note= "V2 region"  
 FT Protein 179..609  
 FT /note= "Surface glycoprotein"  
 FT Region 360..424  
 FT /note= "V3 region"  
 FT Region 451..483  
 FT /note= "V4 region"  
 FT Region 539..566  
 FT /note= "V5 region"  
 FT Protein 610..855  
 FT /note= "Transmembrane protein"  
 PN W09406471-A.  
 PD 31-MAR-1994.  
 PF 20-SEP-1993; G01974.  
 PR 21-SEP-1992; GB-019936.  
 PA (PITM) PITMAN MOORE INC.  
 PI Francis MJ;  
 DR WPI; 94-118168/14.  
 PT Feline immunodeficiency virus antigenic polypeptide(s) and  
 PT nucleic acid - used to prepare prods. for combating or diagnosis  
 PT of feline immunodeficiency infection

PS Disclosure; Figure 2; 68pp; English.  
 CC The sequences given in R51248-R51262 and R58584-85 represent the  
 CC feline immunodeficiency virus (FIV) envelope protein from different  
 CC strains and serotypes. The consensus sequence basedon these, is given  
 CC in R51247. The synthetic FIV peptides of the invention were derived  
 CC principally from a combination of the sequences of the FIV UK 8 and  
 CC Petaluma isolates, a composite sequence of which is given in R51246.  
 CC Variations in the sequence may occur between different strains or  
 CC serotypes, isolates of different geographical origin or even between  
 CC different isolates from the same host.  
 SQ Sequence 855 AA;

Query Match 9.1%; Score 81; DB 10; Length 855;  
 Best Local Similarity 73.3%; Pred. No. 1.98e+01;  
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 789 gilgiglgvllilic 803  
 ||| : ||||| |  
 Qy 31 GLTVILGVLLIGC 45

RESULT 11  
 ID R51248 standard; peptide; 856 AA.  
 AC R51248;  
 DT 08-OCT-1994 (first entry)  
 DE FIV petaluma envelope protein sequence.  
 KW Feline immunodeficiency virus; FIV; vaccine; diagnostic; AIDS;  
 KW T-lymphotropic lentivirus; FIV UK8; FIV Petaluma; envelope protein.  
 OS Synthetic.

FH Key Location/Qualifiers  
 FT Peptide 1..178  
 FT /note= "Hydrophobic leader peptide"  
 FT Region 51..66  
 FT /note= "V1 region"  
 FT Region 95..173  
 FT /note= "V2 region"  
 FT Protein 179..610  
 FT /note= "Surface glycoprotein"  
 FT Region 360..424  
 FT /note= "V3 region"  
 FT Region 451..483  
 FT /note= "V4 region"  
 FT Region 539..567  
 FT /note= "V5 region"  
 FT Protein 611..856  
 FT /note= "Transmembrane protein"  
 PN W09406471-A.  
 PD 31-MAR-1994.  
 PF 20-SEP-1993; G01974.  
 PR 21-SEP-1992; GB-019936.  
 PA (PITM) PITMAN MOORE INC.  
 PI Francis MJ;  
 DR WPI; 94-118168/14.  
 PT Feline immunodeficiency virus antigenic polypeptide(s) and  
 PT nucleic acid - used to prepare prods. for combating or diagnosis  
 PT of feline immunodeficiency infection  
 PS Disclosure; Figure 2; 68pp; English.  
 CC The sequences given in R51248-R51262 and R58584-85 represent the  
 CC feline immunodeficiency virus (FIV) envelope protein from different  
 CC strains and serotypes. The consensus sequence basedon these, is given  
 CC in R51247. The synthetic FIV peptides of the invention were derived  
 CC principally from a combination of the sequences of the FIV UK 8 and  
 CC Petaluma isolates, a composite sequence of which is given in R51246.  
 CC Variations in the sequence may occur between different strains or



CC serotypes, isolates of different geographical origin or even between  
 CC different isolates from the same host.  
 SQ Sequence 856 AA;

Query Match 9.1%; Score 81; DB 10; Length 856;  
 Best Local Similarity 73.3%; Pred. No. 1.98e+01;  
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 790 gilgiglgvlllllc 804  
 ||| : ||||| I  
 Qy 31 GILTVILGVLLIGC 45

RESULT 12  
 ID R51249 standard; peptide; 856 AA.  
 AC R51249;  
 DT 08-OCT-1994 (first entry)  
 DE FIV PET-F14 envelope protein sequence.  
 KW Feline immunodeficiency virus; FIV; vaccine; diagnostic; AIDS;  
 KW T-lymphotropic lentivirus; FIV UK8; FIV Petaluma; envelope protein.  
 OS Synthetic.

FH Key Location/Qualifiers  
 FT Peptide 1..178  
 FT /note= "Hydrophobic leader peptide"  
 FT Region 51..66  
 FT /note= "V1 region"  
 FT Region 95..173  
 FT /note= "V2 region"  
 FT Protein 179..610  
 FT /note= "Surface glycoprotein"  
 FT Region 360..424  
 FT /note= "V3 region"  
 FT Region 451..483  
 FT /note= "V4 region"  
 FT Region 539..567  
 FT /note= "V5 region"  
 FT Protein 611..856  
 FT /note= "Transmembrane protein"  
 PN W09406471-A.  
 PD 31-MAR-1994.  
 PR 20-SEP-1993; G01974.  
 PR 21-SEP-1992; GB-019936.  
 PA (PITM ) PITMAN MOORE INC.  
 PI Francis MJ;  
 DR WPI; 94-118168/14.  
 FT Feline immunodeficiency virus antigenic polypeptide(s) and  
 FT nucleic acid - used to prepare prods. for combating or diagnosis  
 PT of feline immunodeficiency infection  
 PS Disclosure; Figure 2; 68pp; English.  
 CC The sequences given in R51248-R51262 and R58584-85 represent the  
 CC feline immunodeficiency virus (FIV) envelope protein from different  
 CC strains and serotypes. The consensus sequence basedon these, is given  
 CC in R51247. The synthetic FIV peptides of the invention were derived  
 CC principally from a combination of the sequences of the FIV UK 8 and  
 CC Petaluma isolates, a composite sequence of which is given in R51246.  
 CC Variations in the sequence may occur between different strains or  
 CC serotypes, isolates of different geographical origin or even between  
 CC different isolates from the same host.  
 SQ Sequence 856 AA;

Query Match 9.1%; Score 81; DB 10; Length 856;  
 Best Local Similarity 73.3%; Pred. No. 1.98e+01;  
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 790 gilgiglgvlllllc 804  
 ||| : ||||| I  
 Qy 31 GILTVILGVLLIGC 45

RESULT 13  
 ID R51247 standard; peptide; 856 AA.  
 AC R51247;  
 DT 08-OCT-1994 (first entry)  
 DE FIV envelope protein consensus sequence.  
 KW Feline immunodeficiency virus; FIV; vaccine; diagnostic; AIDS;  
 KW T-lymphotropic lentivirus; FIV UK8; FIV Petaluma; envelope protein.  
 OS Synthetic.

FH Key Location/Qualifiers  
 FT Peptide 1..178  
 FT /note= "Hydrophobic leader peptide"  
 FT Region 51..66  
 FT /note= "V1 region"  
 FT Region 95..173  
 FT /note= "V2 region"  
 FT Protein 179..610  
 FT /note= "Surface glycoprotein"  
 FT Region 360..424  
 FT /note= "V3 region"  
 FT Region 451..483  
 FT /note= "V4 region"  
 FT Region 539..567  
 FT /note= "V5 region"  
 FT Protein 611..856  
 FT /note= "Transmembrane protein"  
 PN W09406471-A.  
 PD 31-MAR-1994.  
 PR 20-SEP-1993; G01974.  
 PR 21-SEP-1992; GB-019936.  
 PA (PITM ) PITMAN MOORE INC.  
 PI Francis MJ;  
 DR WPI; 94-118168/14.  
 FT Feline immunodeficiency virus antigenic polypeptide(s) and  
 FT nucleic acid - used to prepare prods. for combating or diagnosis  
 PT of feline immunodeficiency infection  
 PS Disclosure; Figure 2; 68pp; English.  
 CC The sequences given in R51248-R51262 and R58584-85 represent the  
 CC feline immunodeficiency virus (FIV) envelope protein from different  
 CC strains and serotypes. The consensus sequence basedon these, is given  
 CC in R51247. The synthetic FIV peptides of the invention were derived  
 CC principally from a combination of the sequences of the FIV UK 8 and  
 CC Petaluma isolates, a composite sequence of which is given in R51246.  
 CC Variations in the sequence may occur between different strains or  
 CC serotypes, isolates of different geographical origin or even between  
 CC different isolates from the same host.  
 SQ Sequence 856 AA;

Query Match 9.1%; Score 81; DB 10; Length 856;  
 Best Local Similarity 73.3%; Pred. No. 1.98e+01;  
 Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 790 gilgiglgvlllllc 804  
 ||| : ||||| I  
 Qy 31 GILTVILGVLLIGC 45

RESULT 14  
 ID R13844 standard; Protein; 234 AA.  
 AC R13844;





QY 1 MPREDHETGYKKGHSHYTTAEAGIGILTVILGVLLIGCWYRRNGYRALMDK 60

Db 61 slhvotqcaltrrcpoeqfhdhdkvslqekncepvyvnappayeklaeqspppysp 118  
|||||  
QY 61 SLHVGTQALTRCRQEGFHDHDSKVSLEKNCEPVVPNAPPAYEKLSAEQSPPPYSP 118

RESULT 2

ENTRY S01347 #type complete  
TITLE T-cell surface glycoprotein CD2 precursor - mouse  
ALTERNATE\_NAMES CD2 antigen; T11 protein  
ORGANISM #formal name Mus musculus #common name house mouse  
DATE 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 09-Sep-1994

ACCESSIONS S01347

REFERENCE S01347

#authors Clayton, L.K.; Sayre, P.H.; Novotny, J.; Reinherz, E.L.  
#journal Eur. J. Immunol. (1987) 17:1367-1370  
#title Murine and human T11 (CD2) cDNA sequences suggest a common signal transduction mechanism.

#cross-references MUID:88004738

#accession S01347

##molecule\_type mRNA

##residues 1-344 ##label CLA

##cross-references EMBL:X06143

CLASSIFICATION #superfamily T-cell surface glycoprotein CD2  
KEYWORDS glycoprotein; transmembrane protein

FEATURE 1-22

23-344 #domain signal sequence #label SIG\

23-344 #product T-cell surface glycoprotein CD2 #label MAT\

23-203 #domain extracellular #label EXT\

204-228 #domain transmembrane #label TM\

229-344 #domain intracellular #label INT

SUMMARY #length 344 #molecular-weight 38400 #checksum 4741

Query Match

Best Local Similarity 11.1%; Score 99; DB 5; Length 344;

Matches 21; Conservative 9; Mismatches 18; Indels 4; Gaps 4;

Db 198 pekqlsf-yvtvgvqag-qlllavl-valfifc-ickrrknrrrkddee1 245  
| | | : | : | | | | | : | | | : | | | : | | | : | | | :

QY 13 PKKGHSHYTTAEAGIGILTVILGVLLIGCWYRRNGYRALMDKSLHV 64

RESULT 3

ENTRY S02293 #type complete

TITLE T-cell surface glycoprotein CD2 precursor - mouse

ALTERNATE\_NAMES T-lymphocyte antigen CD2

ORGANISM #formal name Mus musculus #common name house mouse

DATE 01-Dec-1989 #sequence\_revision 01-Dec-1989 #text\_change 18-Jun-1993

ACCESSIONS S02293

REFERENCE S02293

#authors Sewell, W.A.; Brown, M.H.; Owen, M.J.; Fink, P.J.; Kozak, C.A.; Crompton, M.J.

#journal Eur. J. Immunol. (1987) 17:1015-1020

#title The murine homologue of the T lymphocyte CD2 antigen: molecular cloning, chromosome assignment and cell surface expression.

#cross-references MUID:87276135

#accession S02293

##status not compared with conceptual translation

##molecule\_type mRNA

##residues 1-344 ##label SEW

#cross-references EMBL:Y00023

GENETICS

#map position 3

CLASSIFICATION #superfamily T-cell surface glycoprotein CD2

KEYWORDS glycoprotein; T-cell

FEATURE 1-22

23-344 #domain signal sequence #label SIG\

23-344 #product T-cell surface glycoprotein CD2 #label MAT

SUMMARY #length 344 #molecular-weight 38414 #checksum 4681

Query Match

Best Local Similarity 11.1%; Score 99; DB 5; Length 344;

Matches 21; Conservative 9; Mismatches 18; Indels 4; Gaps 4;

Db 198 pekqlsf-yvtvgvqag-qlllavl-valfifc-ickrrknrrrkddee1 245  
| | | : | : | | | | | : | | | : | | | : | | | : | | | :

QY 13 PKKGHSHYTTAEAGIGILTVILGVLLIGCWYRRNGYRALMDKSLHV 64

RESULT 4

ENTRY B28967 #type complete

TITLE T-cell surface glycoprotein T11 (CD2) - mouse

ORGANISM #formal name Mus musculus #common name house mouse

DATE 30-Jun-1989 #sequence\_revision 03-Jun-1993 #text\_change 31-Dec-1993

ACCESSIONS B28967

REFERENCE A28967

#authors Diamond, D.J.; Clayton, L.K.; Sayre, P.H.; Reinherz, E.L.

#journal Proc. Natl. Acad. Sci. U.S.A. (1988) 85:1615-1619

#title Exon-intron organization and sequence comparison of human and murine T11 (CD2) genes.

#cross-references MUID:88144486

#accession B28967

##molecule\_type mRNA

##residues 1-344 ##label DIA

##cross-references GB:M19807; GB:J03622; GB:J03623

##note the authors translated the codon TAT for residue 99 as Thr

CLASSIFICATION #superfamily T-cell surface glycoprotein CD2

KEYWORDS glycoprotein

SUMMARY #length 344 #molecular-weight 38325 #checksum 4974

Query Match

Best Local Similarity 11.1%; Score 99; DB 5; Length 344;

Matches 21; Conservative 9; Mismatches 18; Indels 4; Gaps 4;

Db 198 pekqlsf-yvtvgvqag-qlllavl-valfifc-ickrrknrrrkddee1 245  
| | | : | : | | | | | : | | | : | | | : | | | : | | | :

QY 13 PKKGHSHYTTAEAGIGILTVILGVLLIGCWYRRNGYRALMDKSLHV 64

RESULT 5

ENTRY S47272 #type complete

TITLE B-CAM protein - human

ORGANISM #formal name Homo sapiens #common name man

DATE 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 06-Jan-1995

ACCESSIONS S47272

REFERENCE S47272

#authors Campbell, I.G.

#submission submitted to the EMBL Data Library, July 1994

#accession S47272

##status preliminary

##molecule\_type mRNA



<hr/>					
ORGANISM	protein-2.	formal_name Drosophila melanogaster			
DATE		28-May-1992 #sequence_revision 12-Jun-1992 #text_change 25-Aug-1995			
ACCESSIONS	C41214				
REFERENCE	A41214				
#authors	Yang, X.; Seow, K.T.; Bahri, S.M.; Oon, S.H.; Chia, W.				
#journal	Cell (1991) 67:661-673				
#title	Two Drosophila receptor-like tyrosine phosphatase genes are expressed in a subset of developing axons and pioneer neurons in the embryonic CNS.				
#cross-references	MUID:92034988				
#accession	C41214				
##status	preliminary				
##molecule_type	mRNA				
##residues	1-1630 ##label YAN				
##cross-references	GB:M80465				
CLASSIFICATION	#superfamily protein-tyrosine-phosphatase homology				
KEYWORDS	phosphoric monoester hydrolase				
FEATURE	1295-1515	#domain protein-tyrosine-phosphatase homology #label PTP1A			
	1467	#active site Cys (phosphocysteine intermediate) #status predicted\			
	1473	#binding_site substrate phosphate (Arg) #status predicted			
SUMMARY	#length 1630 #molecular-weight 184861 #checksum 8804				
Query Match	9.8%; Score 87; DB 10; Length 1630;				
Best Local Similarity	37.0%; Pred. No. 1.58e+00;				
Matches	17; Conservative 7; Mismatches 19; Indels 3; Gaps 3;				
<hr/>					
Db	1184 ysfplqtd-qdntslvltvp-ltill-vllvtllyfkrnmr 1226				
	:     :             :       :				
Qy	10 YGPKKGCHSYTAEAAAGIGILTVILGVLLLCWCYRRNGR 55				
RESULT	12				
ENTRY	S19243	#type complete			
TITLE	tyrosinase-related protein TRP-2 - mouse				
ORGANISM	#formal_name Mus musculus #common_name house mouse				
DATE	13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 12-Apr-1995				
ACCESSIONS	S19243				
REFERENCE	S19243				
#authors	Jackson, I.J.; Chambers, D.M.; Tsukamoto, K.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.; Hearing, V.				
#journal	EMBO J. (1992) 11:527-535				
#title	A second tyrosinase-related protein, TRP-2, maps to and is mutated at the mouse slaty locus.				
#cross-references	MUID:92164640				
#accession	S19243				
##status	preliminary				
##molecule_type	mRNA				
##residues	1-517 ##label JAC				
##cross-references	EMBL:X63349				
CLASSIFICATION	#superfamily monophenol monooxygenase				
SUMMARY	#length 517 #molecular-weight 58569 #checksum 6717				
Query Match	9.7%; Score 86; DB 4; Length 517;				
Best Local Similarity	41.2%; Pred. No. 2.11e+00;				
Matches	14; Conservative 9; Mismatches 10; Indels 1; Gaps 1;				
<hr/>					
Db	475 igilgafvllglafkgyrrlrkyaplmegtgl 508				
	:     :             :       :				

Qy 30 IGILTVILGVLLIGCWYCR-RNGYRALMDKSL 62

RESULT 13  
ENTRY  
TITLE  
#map\_position 2R  
#domain transmembrane #status predicted #label TM1\  
#domain transmembrane #status predicted #label TM2\  
#domain transmembrane #status predicted #label TM3  
#length 320 #molecular-weight 35986 #checksum 4017  
SUMMARY  
Query Match 9.6%; Score 85; DB 10; Length 320;  
Best Local Similarity 41.7%; Pred. No. 2.82e+00;  
Matches 10; Conservative 8; Mismatches 4; Indels 2; Gaps 2;

Db 287 fgvlsnilvvpfymvcswyirkw 310  
:||: || : :|| || || :  
Qy 30 IGILTVILGV-LLIIGC-WYCRRR 51

RESULT 15  
ENTRY  
TITLE  
#map\_position 2R  
#domain transmembrane #status predicted #label TM1\  
#domain transmembrane #status predicted #label TM2\  
#domain transmembrane #status predicted #label TM3  
#length 320 #molecular-weight 35986 #checksum 4017  
SUMMARY  
Query Match 9.6%; Score 85; DB 10; Length 320;  
Best Local Similarity 41.7%; Pred. No. 2.82e+00;  
Matches 10; Conservative 8; Mismatches 4; Indels 2; Gaps 2;

Db 648 krgetrdvtiavstvtvfavml-iitlsvv-yctrrk-yr 686  
||: : : :||: : : :||: ||  
Qy 14 KKGHSHYTAEEAAGIGILTVILGVLLIGCWYCRRRNGYR 55

RESULT 14  
ENTRY  
TITLE  
#map\_position 2R  
#domain transmembrane #status predicted #label TM1\  
#domain transmembrane #status predicted #label TM2\  
#domain transmembrane #status predicted #label TM3  
#length 320 #molecular-weight 35986 #checksum 4017  
SUMMARY  
Query Match 9.6%; Score 85; DB 10; Length 320;  
Best Local Similarity 41.7%; Pred. No. 2.82e+00;  
Matches 10; Conservative 8; Mismatches 4; Indels 2; Gaps 2;

Db 648 krgetrdvtiavstvtvfavml-iitlsvv-yctrrk-yr 686  
||: : : :||: : : :||: ||  
Qy 14 KKGHSHYTAEEAAGIGILTVILGVLLIGCWYCRRRNGYR 55

RESULT 14  
ENTRY  
TITLE  
#map\_position 2R  
#domain transmembrane #status predicted #label TM1\  
#domain transmembrane #status predicted #label TM2\  
#domain transmembrane #status predicted #label TM3  
#length 320 #molecular-weight 35986 #checksum 4017  
SUMMARY  
Query Match 9.6%; Score 85; DB 10; Length 320;  
Best Local Similarity 41.7%; Pred. No. 2.82e+00;  
Matches 10; Conservative 8; Mismatches 4; Indels 2; Gaps 2;

Qy 30 IGILTVILGV-LLIIGC-WYCRRR 51

Db 287 fgvlsnilvvpfymvcswyirkw 310  
:||: || : :|| || || :  
Qy 30 IGILTVILGV-LLIIGC-WYCRRR 51

RESULT 15  
ENTRY  
TITLE  
#map\_position 2R  
#domain transmembrane #status predicted #label TM1\  
#domain transmembrane #status predicted #label TM2\  
#domain transmembrane #status predicted #label TM3  
#length 320 #molecular-weight 35986 #checksum 4017  
SUMMARY  
Query Match 9.6%; Score 85; DB 10; Length 320;  
Best Local Similarity 41.7%; Pred. No. 2.82e+00;  
Matches 10; Conservative 8; Mismatches 4; Indels 2; Gaps 2;

Db 287 fgvlsnilvvpfymvcswyirkw 310  
:||: || : :|| || || :  
Qy 30 IGILTVILGV-LLIIGC-WYCRRR 51

RESULT 15  
ENTRY  
TITLE  
#map\_position 2R  
#domain transmembrane #status predicted #label TM1\  
#domain transmembrane #status predicted #label TM2\  
#domain transmembrane #status predicted #label TM3  
#length 320 #molecular-weight 35986 #checksum 4017  
SUMMARY  
Query Match 9.6%; Score 85; DB 10; Length 320;  
Best Local Similarity 41.7%; Pred. No. 2.82e+00;  
Matches 10; Conservative 8; Mismatches 4; Indels 2; Gaps 2;

Db 287 fgvlsnilvvpfymvcswyirkw 310  
:||: || : :|| || || :  
Qy 30 IGILTVILGV-LLIIGC-WYCRRR 51

RESULT 15  
ENTRY  
TITLE  
#map\_position 2R  
#domain transmembrane #status predicted #label TM1\  
#domain transmembrane #status predicted #label TM2\  
#domain transmembrane #status predicted #label TM3  
#length 320 #molecular-weight 35986 #checksum 4017  
SUMMARY  
Query Match 9.6%; Score 85; DB 10; Length 320;  
Best Local Similarity 41.7%; Pred. No. 2.82e+00;  
Matches 10; Conservative 8; Mismatches 4; Indels 2; Gaps 2;



Aug 27 07:56

US-08-231-565A-2.ppt

11

Best Local Similarity 34.1%; Pred. No. 2.82e+00;  
Matches 15; Conservative 11; Mismatches 14; Indels 4; Gaps 3;

Db 282 qqgh-ftevlmgigillvlviatlllvwtacfy-rtngivril 323

|:|:| ::|:| ::|:| ::|:| ::|:| ::|:| ::|:|

Qy 16 GHGHSYTTAEAAAGIGILTIVILGVLLI--GCWYRRRNGYRAL 57

Search completed: Tue Aug 27 08:08:31 1996  
Job time : 49 secs.





RA YOKOYAMA K., SUZUKI H., YASUMOTO K.I., TOMITA Y., SHIBAHARA S.;  
 RL BIOCHIM. BIOPHYS. ACTA 1217:317-321(1994).  
 CC -1- CATALYTIC ACTIVITY: DOPACHROME = 5, 6-DIHYDROXYINDOLE-2-CARBOXYLATE.  
 CC -1- COFACTOR: CONTAINS TWO ZINC ATOMS (BY SIMILARITY).  
 CC -1- PATHWAY: MELANIN BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MELANOSOMAL.  
 CC -1- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY.  
 DR EMBL; D17547; D17547.  
 DR PIR; S43510; S43510.  
 DR MIM; 191275; 11TH EDITION.  
 DR PROSITE; PS00497; TYROSINASE 1.  
 DR PROSITE; PS00498; TYROSINASE 2.  
 RW ISOMERASE; ZINC; GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE;  
 KW MELANIN BIOSYNTHESIS.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT CHAIN 24 519 DOPACHROME TAUTOMERASE.  
 FT DOMAIN 24 472 LUMENAL, MELANOSOME (POTENTIAL).  
 FT TRANSMEM 473 493 POTENTIAL.  
 FT DOMAIN 494 519 CYTOPLASMIC (POTENTIAL).  
 FT METAL 189 189 ZINC A (BY SIMILARITY).  
 FT METAL 211 211 ZINC A (BY SIMILARITY).  
 FT METAL 220 220 ZINC A (BY SIMILARITY).  
 FT METAL 369 369 ZINC B (BY SIMILARITY).  
 FT METAL 373 373 ZINC B (BY SIMILARITY).  
 FT METAL 396 396 ZINC B (BY SIMILARITY).  
 FT METAL 170 170 POTENTIAL.  
 FT CARBOHYD 178 178 POTENTIAL.  
 FT CARBOHYD 237 237 POTENTIAL.  
 FT CARBOHYD 300 300 POTENTIAL.  
 FT CARBOHYD 342 342 POTENTIAL.  
 FT CARBOHYD 377 377 POTENTIAL.  
 SQ SEQUENCE 519 AA; 59145 MW; 1422986 CN;  
 Query Match 9.8%; Score 87; DB 8; Length 519;  
 Best Local Similarity 35.3%; Pred. No. 2.41e-01;  
 Matches 12; Conservative 11; Mismatches 10; Indels 1; Gaps 1;  
 Db 477 mgltvalvgllvllafllqrrlrkgytplmethl 510  
 :| |::|::|::| | |::|::| |  
 QY 30 IGLIVILVILLIGCWYCR-RNGYRALDKSL 62  
 RESULT 4  
 ID PTP1 DROME STANDARD; PRT; 1630 AA.  
 AC P35992;  
 DT 01-JUN-1994 (REL. 29, CREATED)  
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)  
 DE PROTEIN-TYROSINE PHOSPHATASE 10D PRECURSOR (EC 3.1.3.48) (RECEPTOR-LINKED PROTEIN-TYROSINE PHOSPHATASE 10D).  
 GN PTP10D.  
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
 OC EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-EMBRYO;  
 RX MEDLINE; 92034989.  
 RA TAN S.-S., TSOUFAS P., ZINN K.;  
 RL CELL 67:675-685(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-EMBRYO;  
 RX MEDLINE; 92034988.

RA YANG X., SEOW K.T., BAHRI S.M., OON S.H., CHIA W.;  
 RL CELL 67:661-673(1991).  
 CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =  
 CC PROTEIN TYROSINE + ORTHOPHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- TISSUE SPECIFICITY: SELECTIVELY EXPRESSED IN A SUBSET OF AXONS AND  
 CC PIONEER NEURONS IN THE EMBRYO.  
 CC -1- SIMILARITY: CONTAINS ONE PROTEIN-TYROSINE PHOSPHATASE DOMAIN AND  
 CC 12 FIBRONECTIN TYPE III-LIKE DOMAINS.  
 CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS THAT DIFFER IN THEIR C-TERMINAL  
 CC TAILS ARE PRODUCED BY ALTERNATIVE SPLICING.  
 DR EMBL; M80538; M80538.  
 DR EMBL; M80465; M80465.  
 DR PIR; C41214; C41214.  
 DR PIR; D41214; D41214.  
 DR PIR; A41215; A41215.  
 DR HSP; P02751; ITT.  
 DR FLYBASE; FBGN0004370; PTP10D.  
 DR PROSITE; PS00383; TYR. PHOSPHATASE.  
 KW TRANSMEMBRANE; HYDROLASE; DUPLICATION; SIGNAL; ALTERNATIVE SPLICING;  
 KW REPEAT.  
 FT SIGNAL 1 34 POTENTIAL.  
 FT CHAIN 35 1630 PROTEIN-TYROSINE PHOSPHATASE 10D.  
 FT DOMAIN 35 1196 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1197 1218 POTENTIAL.  
 FT DOMAIN 1219 1630 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 43 119 FIBRONECTIN TYPE-III.  
 FT DOMAIN 120 214 FIBRONECTIN TYPE-III.  
 FT DOMAIN 215 308 FIBRONECTIN TYPE-III.  
 FT DOMAIN 309 402 FIBRONECTIN TYPE-III.  
 FT DOMAIN 403 493 FIBRONECTIN TYPE-III.  
 FT DOMAIN 494 580 FIBRONECTIN TYPE-III.  
 FT DOMAIN 581 669 FIBRONECTIN TYPE-III.  
 FT DOMAIN 670 766 FIBRONECTIN TYPE-III.  
 FT DOMAIN 767 861 FIBRONECTIN TYPE-III.  
 FT DOMAIN 862 955 FIBRONECTIN TYPE-III.  
 FT DOMAIN 956 1048 FIBRONECTIN TYPE-III.  
 FT DOMAIN 1049 1189 FIBRONECTIN TYPE-III.  
 FT DOMAIN 1290 1532 PROTEIN-TYROSINE PHOSPHATASE.  
 FT CARBOHYD 75 75 POTENTIAL.  
 FT CARBOHYD 106 106 POTENTIAL.  
 FT CARBOHYD 128 128 POTENTIAL.  
 FT CARBOHYD 169 169 POTENTIAL.  
 FT CARBOHYD 212 212 POTENTIAL.  
 FT CARBOHYD 229 229 POTENTIAL.  
 FT CARBOHYD 259 259 POTENTIAL.  
 FT CARBOHYD 289 289 POTENTIAL.  
 FT CARBOHYD 317 317 POTENTIAL.  
 FT CARBOHYD 471 471 POTENTIAL.  
 FT CARBOHYD 486 486 POTENTIAL.  
 FT CARBOHYD 512 512 POTENTIAL.  
 FT CARBOHYD 533 533 POTENTIAL.  
 FT CARBOHYD 588 588 POTENTIAL.  
 FT CARBOHYD 668 668 POTENTIAL.  
 FT CARBOHYD 687 687 POTENTIAL.  
 FT CARBOHYD 719 719 POTENTIAL.  
 FT CARBOHYD 723 723 POTENTIAL.  
 FT CARBOHYD 823 823 POTENTIAL.  
 FT CARBOHYD 841 841 POTENTIAL.  
 FT CARBOHYD 874 874 POTENTIAL.  
 FT CARBOHYD 908 908 POTENTIAL.  
 FT CARBOHYD 925 925 POTENTIAL.  
 FT CARBOHYD 1001 1001 POTENTIAL.  
 FT CARBOHYD 1104 1104 POTENTIAL.

Aug 27 07:55

US-08-231-565A-2.rsp

FT	CARBOHYD	1135	1135	POTENTIAL.
FT	CARBOHYD	1194	1194	POTENTIAL.
FT	ACT_SITE	1467	1467	BY SIMILARITY.
FT	VARSPLIC	1548	1557	GGQVQDLENG -> DDEGIAESGM (IN SHORT FORM).
FT	VARSPLIC	1558	1630	MISSING (IN SHORT FORM).
FT	CONFLICT	1125	1125	Y -> YQ (IN REF. 1).
FT	CONFLICT	1166	1167	IG -> YR (IN REF. 1).
FT	CONFLICT	1172	1172	R -> A (IN REF. 1).
FT	CONFLICT	1457	1457	C -> G (IN REF. 1).
CSQ	SEQUENCE	1630 AA;	184861 MW;	13985973 CN;

```
Query Match      9.8%; Score 87; DB 6; Length 1630;
Best Local Similarity 37.0%; Pred. No. 2.41e-01;
Matches 11; Conservative 7; Mismatches 19; Indels 3; Gaps 3;
```

5	RESULT	TYR2 MOUSE	STANDARD;	PRT;	517 AA.
AC	P298T2;				
AD	01-APR-1993 (REL. 25, CREATED)				
ADT	01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)				
DDT	01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)				
DBE	DOPACHROME TAUTOMERASE PRECURSOR (EC 5.3.3.12) (DT) (DCT) (DOPACHROME				
DE	DELTA-ISOMERASE) (TYROSINASE-RELATED PROTEIN 2) (TRP-2) (TRP2) (SLATY				
DE	LOCUS PROTEIN)				
DN	TYR2 OR TYRP-2.				
GN	MUS MUSCULUS (MOUSE) .				
OS	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; RODENTIA.				
OC	[1]				
RP	SEQUENCE FROM N.A.				
RP	MEDLINE; 92164640.				
EX	JACKSON I.-J., CHAMBERS D.M., TSUKAMOTO K., COPELAND N.G.,				
RA	GILBERT D.J., JENKINS N.A., HEARING V.J.;				
RA	EMBO J. 11:527-536(1992).				
RL	[2]				
RL	CHARACTERIZATION.				
RP	MEDLINE; 92164639.				
EX	TSUKAMOTO K., JACKSON I.J., URABE K., MONTAGUE P.M., HEARING V.J.;				
RL	EMBO J. 11:519-526(1992).				

1-3] ZINC-BINDING.  
MEDLINE; 95071460.  
X R SOLANO F., MARTINEZ-LIARTE J.H., JIMENEZ-CERVANTES C.,  
X R GARCIA-BORRON J.C., LOZANO J.A.;  
X R BIOCHEM. BIOPHYS. RES. COMMON. 204:1243-1250 (1994).  
X R [4]  
X R ZINC-BINDING.  
X R SOLANO F., JIMENEZ-CERVANTES C., MARTINEZ-LIARTE J.H.,  
X R GARCIA-BORRON J.C., JARA J.R., LOZANO J.A.;  
X R BIOCHEM. J. 0:0-0 (1996).  
X R -1- CATALYTIC ACTIVITY: DOPACHROME = 5, 6-DIHYDROXYINDOLE-2-CARBOXYLATE.  
X R -1- COFACTOR: CONTAINS TWO ZINC ATOMS.  
X R -1- PATHWAY: MELANIN BIOSYNTHESIS.  
X R -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MELANOSOMAL.  
X R -1- DISEASE: DEFECTS IN TYRP2 LEADS TO A DECREASE OF DT ACTIVITY.  
X R -1- AND A CONSEQUENT CHANGE IN THE PIGMENTATION OF THE MICE TO A DARK GREY/BROWN EUMELANIN.  
X R -1- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY.

Query Match 9.6%; Score 85; DB 9; Length 320;  
Best Local Similarity 41.7%; Pred. No. 4.82e-01.

Aug 27 07:55

US-08-231-565A-2.fsp

DR	EMBL; X63349; X63349.
DR	PIR; S19243; S19243.
DR	PROSITE; PS00497; TYROSINASE 1.
DR	PROSITE; PS00498; TYROSINASE 2.
KW	ISOMERASE; ZINC; GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE;
FT	MELANIN BIOSYNTHESIS; DISEASE MUTATION.
FT	SIGNAL
FT	1 23 POTENTIAL.
FT	CHAIN 24 517 DOPACHROME TAUTOMERASE.
FT	DOMAIN 24 472 LUMENAL MELANOSOME (POTENTIAL).
FT	TRANSMEM 473 491 POTENTIAL.
FT	DOMAIN 492 517 CYTOPLASMIC (POTENTIAL).
FT	METAL 189 189 ZINC A (BY SIMILARITY).
FT	METAL 211 211 ZINC A (BY SIMILARITY).
FT	METAL 220 220 ZINC A (BY SIMILARITY).
FT	METAL 369 369 ZINC B (BY SIMILARITY).
FT	METAL 373 373 ZINC B (BY SIMILARITY).
FT	METAL 396 396 ZINC B (BY SIMILARITY).
FT	CARBOHYD 92 92 *POTENTIAL.
FT	CARBOHYD 170 170 POTENTIAL.
FT	CARBOHYD 178 178 POTENTIAL.
FT	CARBOHYD 237 237 POTENTIAL.
FT	CARBOHYD 300 300 POTENTIAL.
FT	CARBOHYD 342 342 POTENTIAL.
FT	CARBOHYD 377 377 POTENTIAL.
FT	VARIANT 194 194 R -> Q (IN SLATY).
SO	SEQUENCE 517 AA: 58569 MW: 1366219 GN:

```

Query Match      9.7%; Score 86; DB 8; Length 517;
Best Local Similarity 41.2%; Pred. No. 3.41e-01;
Matches 14; Conservative 9; Mismatches 10; Indels 1; Gaps 1;

Ddb 475 iglgafvllllglaflcyrrlrkyaplmgtgl 508
      |||| : : || : : || : || : || : :
Ov 30 IGLFVLIVGLLLIGCWYCR--NGYRALMDKSL 62

```

6	RESULT	YB9K YEAST	STANDARD;	PRT;	320 AA.
AC	P38342;				
DT	01-OCT-1994	(REL. 30, CREATED)			
DT	01-OCT-1994	(REL. 30, LAST SEQUENCE UPDATE)			
DT	01-OCT-1994	(REL. 30, LAST ANNOTATION UPDATE)			
DE	HYPOTHETICAL 36.0 KD PROTEIN IN SHM1-WRPL37 INTERGENIC REGION.				
GN	YB265W OR YB1734.				
GN	SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).				
OC	EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.				
RI	[1]				
RR	SEQUENCE FROM N.A.				
RP	STRAIN=S288C;				
RC	MEDLINE: 9320397.				
RX	DOIGNON F., BITEAU N., CROUZET M., AIGLE M.;				
RA	YEAST 9:189-199(1993).				
RL	-!- SURCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL)				
RC					
OR	EMBL; X70529; X70529.				
OR	EMBL; Z36134; Z36134.				
OR	PIR; S32966; S32966.				
OR	HYPOTHETICAL PROTEIN; TRANSMEMBRANE.				
FW	TRANSMEM 162 182				
FT	POTENTIAL.				
FT	TRANSMEM 255 275				
FT	POTENTIAL.				
FT	TRANSMEM 280 300				
FT	POTENTIAL.				
ESQ	SEQUENCE 320 AA; 35986 MW; 532496 CN;				

Query Match 9.6%; Score 85; DB 9; Length 320;  
 Best Local Similarity 41.7%; Pred. No. 4.82e-01;

Matches 10; Conservative 8; Mismatches 4; Indels 2; Gaps 2;

Db 287 fgvlslvlpfymgscwyrkw 310

Qy 30 IGLTWILGW-LLIIGC-WYCRRR 51

RESULT 7  
ID ATH1 YEAST STANDARD; PRT; 918 AA.

AC P05030;

DT 13-AUG-1987 (REL. 05, CREATED)

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)

DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)

DE PLASMA MEMBRANE ATPASE 1 (EC 3.6.1.35) (PROTON PUMP).

GN PMA1.

OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).

OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 86146844.

RA SERRANO R., KIELLAND-BRANDT M.C., FINK G.R.;

RL NATURE 319:689-693(1986).

CC -/- FUNCTION: THE PLASMA MEMBRANE ATPASE OF PLANTS AND FUNGI IS A

CC HYDROGEN ION PUMP. THE PROTON GRADIENT IT GENERATES DRIVES THE

CC ACTIVE TRANSPORT OF NUTRIENTS BY H+ SYMPORT. THE RESULTING

CC EXTERNAL ACIDIFICATION AND/OR INTERNAL ALKINIZATION MAY MEDIATE

CC GROWTH RESPONSES.

CC -/- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + ORTHOPHOSPHATE.

CC -/- THERE ARE TWO PLASMA MEMBRANE ATPASES IN YEAST. THIS IS THE

CC MAJOR ISOFORM.

CC -/- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -/- PTM: PHOSPHORYLATED ON MULTIPLE SER AND THR RESIDUES.

CC -/- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY

CC (EI-E2 ATPASES).

DR EMBL; X03534; X03534.

DR PIR; A25823; P8BY1P.

DR LISTA; SC00832; PMA1.

DR SGD; L0001449; PMA1.

DR PROSITE; PS00154; ATPASE\_E1\_E2.

KW HYDROLASE; HYDROGEN ION TRANSPORT; TRANSMEMBRANE; PHOSPHORYLATION;

KW ATP-BINDING; MULTIGENE FAMILY; GLYCOPROTEIN.

FT DOMAIN 1 115

FT TRANSMEM 116 136

FT DOMAIN 137 140

FT TRANSMEM 141 160

FT DOMAIN 161 291

FT TRANSMEM 292 313

FT DOMAIN 314 325

FT TRANSMEM 325 347

FT DOMAIN 348 719

FT TRANSMEM 720 738

FT DOMAIN 739 754

FT TRANSMEM 755 774

FT DOMAIN 775 824

FT TRANSMEM 825 845

FT DOMAIN 846 857

FT TRANSMEM 858 874

FT DOMAIN 875 918

FT MOD\_RES 378 378

FT BINDING 474 474

FT CARBOHYD 848 848

FT DOMAIN 15 25

FT DOMAIN 31 78

SQ SEQUENCE 918 AA; 99619 MW; 4420147 CN;

Query Match 9.6%; Score 85; DB 1; Length 918;

Best Local Similarity 34.1%; Pred. No. 4.82e-01;

Matches 15; Conservative 11; Mismatches 14; Indels 4; Gaps 3;

Db 282 gqgh-ftevlmgigillvltatllvwtacfy-rtnqivrrll 323

Qy 16 GHGSYTTAEAGIGILTIVILGVLLLI--GWYCRRRNGYRAL 57

RESULT 8

ID CDIA HUMAN STANDARD; PRT; 327 AA.

AC P06126;

DT 01-JAN-1988 (REL. 06, CREATED)

DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)

DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)

DE T-CELL SURFACE GLYCOPROTEIN CDIA PRECURSOR (CDIA ANTIGEN) (T-CELL

DE SURFACE ANTIGEN T6/LEU-6) (HTAI THYMOCYTE ANTIGEN).

GN CDIA.

OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

OC EUTHERIA; PRIMATES.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 88097453.

RA MARTIN L.H., CALABI F., LEFEBVRE F.-A., BILSLAND C.A.G., MILSTEIN C.;

RL PROC. NATL. ACAD. SCI. U.S.A. 84:9189-9193(1987).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE; 89341413.

RA ARUFFO A., SEED B.;

RL J. IMMUNOL. 143:1723-1730(1989).

RN [3]

RP SEQUENCE OF 99-327 FROM N.A.

RC TISSUE=T-CELL;

RX MEDLINE; 87014824.

RA CALABI F., MILSTEIN C.;

RL NATURE 323:540-543(1986).

CC -/- FUNCTION: NOT KNOWN.

CC -/- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC -/- SUBUNIT: ASSOCIATES NON-COVALENTLY WITH BETA-2-MICROGLOBULIN.

CC -/- TISSUE SPECIFICITY: EXPRESSED ON CORTICAL THYMOCYTES, ON CERTAIN

CC T-CELL LEUKEMIAS, AND IN VARIOUS OTHER TISSUES.

CC -/- SIMILARITY: TO OTHER CD1 ANTIGENS, AND TO MHC CLASS I ANTIGENS.

CC -/- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.

DR EMBL; M28825; M28825.

DR EMBL; M22080; M22080.

DR EMBL; M22163; M22163.

DR EMBL; M22164; M22164.

DR EMBL; M22165; M22165.

DR EMBL; M22166; M22166.

DR EMBL; M22167; M22167.

DR EMBL; X04450; X04450.

DR PIR; A02242; HLHUCD.

DR PIR; A39957; A39957.

DR PIR; A45801; A45801.

DR HSP; P01842; IDLH.

DR MM; 188370; 11TH EDITION.

KW GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE; IMMUNOGLOBULIN FOLD; T-CELL;

KW MULTIGENE FAMILY.

FT SIGNAL 1 16

FT CHAIN 17 327

FT DOMAIN 17 108

FT DOMAIN 109 201

CDIA ANTIGEN.

EXTRACELLULAR ALPHA-1.

EXTRACELLULAR ALPHA-2.

Aug 27 07:55

US-08-231-565A-2.rsp

11

FT DOMAIN 202 294 EXTRACELLULAR ALPHA-3.  
FT TRANSMEM 300 321 POTENTIAL.  
FT DOMAIN 322 327 CYTOPLASMIC (POTENTIAL).  
FT DISULFID 119 183 BY SIMILARITY.  
FT CARBOHYD 37 37 POTENTIAL.  
FT CARBOHYD 60 60 POTENTIAL.  
FT CARBOHYD 74 74 POTENTIAL.  
FT CARBOHYD 145 145 POTENTIAL.  
FT CONFLICT 30 30 I -> T (IN REF. 1).  
FT CONFLICT 68 68 W -> C (IN REF. 1).  
SQ SEQUENCE 327 AA; 37172 MW; 559220 CN;

Query Match 9.4%; Score 84; DB 2; Length 327;  
Best Local Similarity 48.1%; Pred. No. 6.78e-01;  
Matches 13; Conservative 7; Mismatches 4; Indels 3; Gaps 3;

Db 299 svqfilavlpvlllglalwf-kr 324  
QY : | : ||:| : ||||| | : | : |  
27 AAGIGILTILGVLGCG-C-WYCR 51

RESULT 9

ID CD80 MOUSE STANDARD; PRT; 306 AA.  
AC Q00609;  
DT 01-FEB-1994 (REL. 28, CREATED)  
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE T LYMPHOCYTE ACTIVATION ANTIGEN CD80 PRECURSOR (ACTIVATION B7-1  
DE ANTIGEN) (B7).  
GN CD80 OR B7.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; RODENTIA.  
RN SEQUENCE FROM N.A.  
RC TISSUE=B-CELL;  
RX MEDLINE; 91341422.  
RA GRAY G.S., FREEMAN G.J., GIMMI C.D., LOWARD D.B., ZHOU L.J.,  
RA WHITE M., FINGEROTH J.D., GRIBBEN J.G., NADLER L.M.;  
RL J. EXP. MED. 174:625-631 (1991).  
RN [2]

RP SEQUENCE FROM N.A.  
RC TISSUE=B-CELL;  
RX MEDLINE; 93307789.  
RA SELVAKUMAR A., WHITE P.C., DUPONT B.;  
RL IMMUNOGENETICS 38:292-295 (1993).  
CC -!- FUNCTION: INVOLVED IN THE COSTIMULATORY SIGNAL ESSENTIAL FOR T  
CC LYMPHOCYTES ACTIVATION. T CELL PROLIFERATION AND CYTOKINE  
CC PRODUCTION IS INDUCED BY THE BINDING OF CD28 OR CTLA-4 TO THIS  
CC RECEPTOR.  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -!- DEVELOPMENTAL STAGE: EXPRESSED BETWEEN 4 AND 12 HOURS POST-  
CC ACTIVATION. PROTEIN WAS DETECTED AT CELL SURFACE AT 24 HOURS AND  
CC IT'S EXPRESSION WAS MAXIMAL FROM 48 TO 72 HOURS POST-ACTIVATION.  
CC -!- TISSUE SPECIFICITY: EXPRESSED ON ACTIVATED B CELLS, GAMMA  
CC INTERFERON STIMULATED MONOCYTES AND NONCIRCULATING B-CELL  
CC MALIGNANCIES.  
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS  
CC ONE C2-LIKE AND ONE V-LIKE DOMAIN.  
DR EMBL; X60958; X60958.  
DR EMBL; L12587; L12585.  
DR EMBL; L12586; L12586.  
DR EMBL; L12587; L12587.  
DR EMBL; L12588; L12588.

Aug 27 07:55

US-08-231-565A-2.rsp

12

DR EMBL; L12589; L12589.  
DR PIR; S17291; S17291.  
KW IMMUNOGLOBULIN FOLD; T-CELL; GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE.  
FT SIGNAL 1 37  
FT CHAIN 38 306  
FT DOMAIN 22 246  
FT TRANSMEM 247 268  
FT DOMAIN 269 306  
FT DOMAIN 47 126  
FT DOMAIN 158 226  
FT DOMAIN 227 246  
FT DISULFID 54 119  
FT DISULFID 165 219  
FT CARBOHYD 93 93  
FT CARBOHYD 99 99  
FT CARBOHYD 149 149  
FT CARBOHYD 189 189  
FT CARBOHYD 210 210  
FT CARBOHYD 214 214  
SQ SEQUENCE 306 AA; 34589 MW; 505533 CN;

Query Match 9.3%; Score 83; DB 2; Length 306;  
Best Local Similarity 27.6%; Pred. No. 9.52e-01;  
Matches 8; Conservative 14; Mismatches 6; Indels 1; Gaps 1;

Db 252 gagfgavltvvvlvilkc-fchrsr 279  
QY : ||:| : : : ||:| : ||:| : |  
27 AAGIGILTILGVLGCG-CYCRNGYR 55

RESULT 10

ID INGR MOUSE STANDARD; PRT; 477 AA.  
AC P15261;  
DT 01-APR-1990 (REL. 14, CREATED)  
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE INTERFERON-GAMMA RECEPTOR ALPHA CHAIN PRECURSOR.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; RODENTIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 90083245.  
RA MUNRO S., MANIATIS T.;  
RL PROC. NATL. ACAD. SCI. U.S.A. 86:9248-9252 (1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 90046824.  
RA GRAY P.W., LEONG S., FENNIE E.H., FARRAR M.A., PINGEL J.T.,  
RA FERNANDEZ-LUNA J., SCHREIBER R.D.;  
RL PROC. NATL. ACAD. SCI. U.S.A. 86:8497-8501 (1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 90099370.  
RA HEMMI S., PEGHINI P., METZLER M., MERLIN G., DEMBIC Z., AGUET M.;  
RL PROC. NATL. ACAD. SCI. U.S.A. 86:9901-9905 (1989).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 90036866.  
RA KUMAR C.S., MUTHUKUMARAN G., FROST L.J., NOE M., AHN Y.H.,  
RA MARIANO T.M., PESTKA S.;  
RL J. BIOL. CHEM. 264:17939-17946 (1989).  
RN [5]  
RP SEQUENCE FROM N.A.

RESULT	11	STANDARD;	PRT;	897 AA.
ID	CYRB HUMAN			
AC	P3297;			
DT	01-OCT-1993 (REL. 27, CREATED)			
DT	01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)			
DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)			
DE	CTOKINE RECEPTOR COMMON BETA CHAIN PRECURSOR.			
EN	CSF2RB OR IL5RB OR IL3RB.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; PRIMATES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 91089571.			
RA	HAYASHIDA K., KITAMURA T., GORMAN D.M., ARAI K., YOKOTA T.,			

FI	TRANSMEM	789	806	POTENTIAL
SQ	SEQUENCE	1010 AA;	111678 MW;	5433252 CN;

Db 528 gfgdsevsplrtiedphvcddpspgpdttpaasdlpteqpssp 569  
|| : | :::: : | : || : || : || : |  
Ov 78 GEDHRD-SKVSLOE-KN-CEPVV-PNAPPAVEKLSAQSPPP 115



Query Match 9.1%; Score 81; DB 3; Length 511;  
Best Local Similarity 36.7%; Pred. No. 1.86e+00;  
Matches 11; Conservative 9; Mismatches 7; Indels 3; Gaps 3;

17 HCHSYTTAEERAGIGILTVILGVLLIGCW 46

RESULT	14	
D	MCP2 SALTY	STANDARD; PRT; 553 AA.
C	P02941;	
T	21-JUL-1986	(REL. 01, CREATED)
T	01-NOV-1995	(REL. 32, LAST SEQUENCE UPDATE)
T	01-NOV-1995	(REL. 32, LAST ANNOTATION UPDATE)
E	METHYL-ACCEPTING CHEMOTAXIS PROTEIN II (MCP-II) (ASPARTATE	
E	CHEMORECEPTOR PROTEIN).	
N	TAR.	
N	SALMONELLA TYPHIMURUM.	
S	PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;	
C	ENTEROBACTERIACEAE.	
C	[1]	
N	SEQUENCE FROM N. A.	
P	MEDLINE; 83197387.	
X	A RUSSO A.F., KOSHLAND D.E. JR.;	
L	SCIENCE 220:1016-1020(1983).	
N	[2]	
N	REVISIONS TO 236 AND INSERTION OF ALA-498.	
P	STRAIN=LT2;	
C		
A	STOCK A.;	
L	THESIS (1986), UNIVERSITY OF BERKELEY, U.S.A.	
L	[3]	

YEH J.I., BIEMANN H.-P., PANDIT J., KOSHLAND D.E., KIM S.-H.;  
J. BIOL. CHEM. 268:9787-9792(1993).  
-I- FUNCTION: RECEPTOR FOR THE ATTRACTANT L-ASPARTATE AND RELATED AMINO AND DICARBOXYLIC ACIDS. TAR MEDIATES TAXIS AWAY FROM THE REPELLENTS COBALT AND NICKEL. UNLIKE E.COLI TAR IT DOES NOT MEDIATES MALTOSE TAXIS.  
-I- FUNCTION: CHEMOTACTIC-SIGNAL TRANDCERS RESPOND TO CHANGES IN THE CONCENTRATION OF ATTRACTANTS AND REPELLENTS IN THE ENVIRONMENT, AND TRANSDUCE A SIGNAL FROM THE OUTSIDE TO THE INSIDE OF THE CELL, AND FACILITATE SENSORY ADAPTATION THROUGH THE VARIATION OF THE LEVEL OF METHYLATION. ATTRACTANTS INCREASE THE LEVEL OF METHYLATION WHILE REPELLENTS DECREASE THE LEVEL OF METHYLATION, THE METHYL GROUPS ARE ADDED BY THE METHYLTRANSFERASE CHER AND REMOVED BY THE METHYLESTERASE CHER.  
-I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.  
-I- SIMILARITY: TO THE OTHER BACTERIAL CHEMOTAXIS SENSORY TRANSDUCERS. EMBL: J01809; J01809.  
PIR: A03441; QREBDT.  
PDB: 1WAS; 15-JAN-95.  
PDB: 1WAT; 15-JAN-95.  
PIR: STYGENE; SG10387; TAR.  
PROSITE: PS00538; CHEMOTAXIS TRANSDUCER.  
CHEMOTAXIS; TRANSMEMBRANE; METHYLATION; PERIPLASMIC; OUTER MEMBRANE; 3D-STRUCTURE.  
DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).  
TRANSMEM 7 33 POTENTIAL.  
CHAIN 34 552 TAR PROTEIN.  
DOMAIN 34 190 PERIPLASMIC (POTENTIAL).  
TRANSMEM 191 215 POTENTIAL.  
DOMAIN 216 552 CYTOPLASMIC (POTENTIAL).  
DOMAIN 64 73 THE 3 ARG MAY FORM A POSITIVELY CHARGED

I	DOMAIN	ZIG	JJZ	CYCLOELASTIC (POTENTIAL):
T	DOMAIN	64	73	THE 3 ARG MAY FORM A POSITIVELY CHARGED

Aug 27 07:55

US-08-231-565A-2.rsp

17

POCKET, WHICH BINDS THE ALPHA-CARBOXYL

GROUP OF THE ATTRACTANT AA.

DEMAMINATION AND METHYLATION.

METHYLATION.

DEMAMINATION AND METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

METHYLATION.

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

(MIL)

Result	No.	Score	Query		DB	ID	Description	Pred. No.
			Match	Length				
	1	1392	98.0	1559	17	T02714	MART-1 melanoma anti	1.35e+08
	2	1372	96.6	354	12	Q76370	Tumour rejection anti	8.83e-107
	3	1332	93.8	760	17	T11550	Tumour rejection anti	3.77e-103
	4	443	31.2	1358	17	T11549	Tumour rejection anti	2.85e-24
c	5	159	11.2	416	4	N10053	Generic coding sequen	9.55e-02
c	6	157	11.1	231	2	N60394	Desulphatohirudin (I)	1.31e-01
	7	138	9.7	1467	3	Q20927	Leptospira hardjo-bov	2.47e+00
	8	132	9.3	3556	11	O63952	VEF gene.	6.04e+00



FT /note="This region is 4.7-5.3 kilobases in length  
FT but the sequence of this region has not been  
FT deduced."

```
FT /note= "encodes Leu"
FT unsure 38..40
FT /*tag= e
```

FT /note="This region is 4.7-5.3 kilobases in length  
FT but the sequence of this region has not been  
FT deduced."  
FT



The sequence was obtained by shearing *L. hardjo-bovis* total genomic DNA to 5 kbp fragments. These were denatured to single strands and allowed to reanneal slowly to form fragments containing the double stranded sequence of 1.4 kbp flanked by shorter regions of single stranded DNA. These regions were digested away using S1 nuclease and the digest product run on a standard agarose gel. The isolated DNA fragment was inserted into the pUC18 cloning vector using T4 DNA ligase. The sequence is a repetitive element present 40 or more times in the *L. hardjo-bovis* genome. Therefore this element provides

FT	/product= Virus enhancing factor
PN	EP-596508-A.
PD	11-MAY-1994.
PF	04-NOV-1993; 117929
PR	04-NOV-1992; US-971624.
PPA	(BOYC-) BOYCE THOMPSON INST PLANT RES.
PI	Granados RR, Hashimoto Y;
PI	WFI; 94-152908/19.
DR	P-FSDB; R53963.
DR	
PT	Gene encoding protein which enhances viral infection of host
PT	insects - for use in novel viral pesticides with increased
PT	efficacy, or for construction of transgenic plants with increased
PT	protection against insects
PT	Claim 3; Fig 3; 32pp; English.
CC	This sequence encodes a virus enhancing factor protein (VEF) which
CC	may also be referred to as the synergistic factor (SF). This
CC	sequence was isolated from Trichoplusia ni (cabbage looper)
CC	granulosis virus (TrGV) and the VEF protein causes rapid degradation
CC	of the petrirophic membrane which lines the midgut lumen of









17

Db 114 ngarcarcargcngaygaygaraaywsngcngaymgncny 154  
| : | : | : : : : | : : | : :  
cp 293 GGNACNAGGGYTCRCARTTYYTCYCTGNARNWSNACYTT 253

Search completed: Tue Aug 27 18:42:40 1996  
Job time : 44 secs.

\*\*\*\*\*

AVPQSRLEA (TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_tpn n.a. - n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Tue Aug 27 18:43:02 1996; MasPar time 164.99 Seconds  
Tabular output not generated.

Title: >US-08-231-565A-2

Description: (1-118) from US08231565A.pep

Perfect Score: 1420

N.A. Sequence: 1 ATGCNCNMGARGAYGCNCA.....SNCCNCCNCTAVNSNCN 354  
Comp: TAGCCNKCNTCTRCNGT.....WNGGNGGNGNATRSWNGN

Scoring table:

Gap 30

Nmatch STD : Dbase 0; Query 0

Searched: 509049 seqs, 177673129 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: EST-STS  
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8  
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14  
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20  
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26  
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32  
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38  
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44  
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50  
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56  
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62  
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68  
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74  
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80  
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86  
87:EST87 88:EST88 89:EST89 90:EST90 91:STS1 92:STS2  
93:STS3 94:STS4 95:STS5 96:STS6 97:STS7  
EST-STS-TWO  
98:EST91 99:EST92 100:EST93 101:EST94 102:EST95  
103:EST96 104:EST97 105:EST98 106:EST99 107:EST100  
108:EST101 109:EST110:gnEST2 111:gnEST3 112:gnEST4  
113:gnEST5 114:gnEST6 115:gnEST7 116:gnEST8 117:gnEST9  
118:gnEST10 119:gnEST11 120:gnEST12 121:gnEST13  
122:gnEST14 123:gnEST15 124:gnSTS1 125:gnSTS2 126:enEST1

Database:

127:enEST2 128:enEST3 129:enEST4 130:enEST5 131:enEST6  
132:enEST7 133:enEST8 134:enEST9 135:enEST10 136:enEST11

Statistics: Mean 61.375; Variance 71.660; scale 0.856

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	1392	98.0	593	113	N23770	yx35a12.rl Homo sapie	0.00e+00
2	1392	98.0	593	136	HS770262	yx35a12.rl Homo sapie	0.00e+00
3	1271	89.5	370	117	N35680	yx62h05.rl Homo sapie	0.00e+00
4	1271	89.5	370	135	HS680270	yx62h05.rl Homo sapie	0.00e+00
5	1205	84.9	393	135	HS656271	yx91a09.rl Homo sapie	0.00e+00
6	1205	84.9	393	117	N36656	yx91a09.rl Homo sapie	0.00e+00
7	947	66.7	303	82	T27884	EST19370 Homo sapiens	4.34e-264
8	722	50.8	473	129	HS091255	yx36a12.sl Homo sapie	1.83e-186
9	722	50.8	473	112	N20091	yx36a12.sl Homo sapie	1.83e-186
10	640	45.1	433	135	HS677260	yx91a09.sl Homo sapie	1.17e-158
11	640	45.1	433	113	N24677	yx91a09.sl Homo sapie	1.17e-158
12	540	38.0	396	136	HS796259	yx97f10.sl Homo sapie	2.94e-125
13	540	38.0	396	113	N24796	yx97f10.sl Homo sapie	2.94e-125
14	530	37.3	401	134	HS562267	yx62h05.sl Homo sapie	5.89e-122
15	530	37.3	401	114	N26562	yx62h05.sl Homo sapie	5.89e-122
16	153	10.8	312	129	HS037250	yy15c06.sl Homo sapie	2.17e-09
17	153	10.8	312	109	H94037	16649 Arabidopsis tha	2.17e-09
18	153	10.8	515	72	R90294	yx97f10.rl Homo sapie	1.14e-05
19	136	9.6	43	117	N36215	yx97f10.rl Homo sapie	1.14e-05
20	136	9.6	43	131	HS215288	ye73a04.rl Homo sapie	7.70e-05
21	132	9.3	411	45	R00378	yl16c08.rl Homo sapie	1.97e-04
22	130	9.2	374	12	H27083	yl15e08.rl Homo sapie	1.97e-04
23	130	9.2	394	12	H26878	yy08d01.rl Homo sapie	1.97e-04
24	130	9.2	479	119	N42998	yu75e03.rl Homo sapie	1.97e-04
25	130	9.2	484	28	H80481	yl16a08.rl Homo sapie	1.97e-04
26	130	9.2	550	12	H27071	18581 Arabidopsis tha	4.97e-04
27	128	9.0	572	126	AT35416	18581 Arabidopsis tha	4.97e-04
28	128	9.0	572	117	N37354	H. sapiens partial cd	7.86e-04
29	127	8.9	289	33	HSB89F102	H. sapiens partial cd	7.86e-04
30	127	8.9	293	32	HHEA61P	H. sapiens partial cd	1.24e-03
31	126	8.9	338	33	HSBA2A042	H. sapiens partial cd	1.24e-03
32	126	8.9	393	62	RS9572	yg97h10.sl Homo sapie	1.24e-03
33	125	8.8	434	9	H18419	yn49g10.sl Homo sapie	1.95e-03
34	123	8.7	307	42	HUM13601B	Human fetal brain cdn	4.77e-03
35	122	8.6	220	73	R95571	SMNHADA008030SK Schis	7.43e-03
36	122	8.6	274	26	H74486	446 Brassica napus cd	7.43e-03
37	122	8.6	289	38	HSC2PH091	H. sapiens partial cd	7.43e-03
38	122	8.6	300	77	T00011	wEST00732 Caenorhabdi	7.43e-03
39	122	8.6	313	32	HHEF052	H. sapiens partial cd	7.43e-03
40	122	8.6	320	73	R95547	SMNHADA007019T3 Schis	7.43e-03
41	122	8.6	343	22	H62105	yu40f05.rl Homo sapie	7.43e-03
42	122	8.6	396	107	T96010	ye47q01.rl Homo sapie	7.43e-03
43	122	8.6	423	22	H61302	yu39a03.rl Homo sapie	7.43e-03
44	122	8.6	442	107	T95596	ye40b03.sl Homo sapie	7.43e-03
45	122	8.6	528	89	T55327	yb47h10.sl Homo sapie	7.43e-03

ALIGNMENTS

RESULT 1





```

FT /organism="Homo sapiens"
FT /clones="266361"
FT /note="human"
FT mRNA
FT <1..>370
FT Sequence 370 BP; 103 A; 87 C; 94 G; 86 T; 0 other;
FT SQ

```

Query Match 89.5%; Score 1271; DB 135; Length 370;  
Best Local Similarity 58.6%; Pred. No. 0.00e+00;  
Matches 197; Conservative 78; Mismatches 59; Indels 2; Gaps 2;

36 atgccagaagaatgctcacttcattcatctatggtaccccaagaaggcgacggccactct 95  
||||| : |||:|:| |||:|:|:| |||:|:|:| |||:|:|:| |||:|:|:| |||:|:|:|

1 ATGCNMGNGARGAYGCNCAYTTATHTAYGNTAYCNAARAARGNCAYGNCAYNSN 60

96 tacaccacggcctgaagagccgcgtgggatacggcatacctgacagtgatcctgggagcttta 155  
||:|| || || || || || || || || || || || || || || || || || || || || ||  
61 TATACNACGNCGAR GARGCGNCGNATHCGCNATHYTNACNGTWTATHYNGGNGNYTN 120

[illegible][illegible]

Db 275 catcggacagcaaaagtgtctcttcaagagaaaaactgtgaacctgtggttcccaatgct 334  
|||:||:||:::||:||:||:||:||:||:||:||:||:||:||:||:||

Db 335 ccacctgcttattggagaaactctctgcgaacagtcc 370  
|| || || || || : || :: | : || :: || ::

RESULT 5

AC N3656;  
DT 20-JAN-1996 (Rel. 46, Created)  
DT 20-JAN-1996 (Rel. 46, Last updated, Version 1)

[illegible]

OC	Eukaryota; Mammalia; Vertebrata; Mammalia;
OC	Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
OC	[1]
RP	1-393

RA Hillier L., Clark N., Dubouque I., Elliston K., Hawkins M.,  
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,  
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,  
RA Trevisan E., Watson D., Willingham A., Willingham D., Willson D.

RL Unpublished.  
CC Contact: Wilson RK WashU-Merck EST Project Washington University  
CC "The WashU-Merck EST Project";  
CC School of Medicine 4444 Parkview Dr. 6501 St. Louis

CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu High quality sequence stops: 169 Source: IMAGE Consortium, LNL This clone is available royalty-free through LNL  
CC  
CC  
CC

CC	information. NCBI gi: 1157798
FH	Key
FH	Location/Qualifiers
FH	

```

FT      /organism="Homo sapiens"
FT      /clone="269080"
FT      /note="human"
FT      mRNA
SQ      <1..>393
SQ      Sequence 393 BP; 105 A; 94 C; 105 G; 81 T; 8 other;

```

Query Match	84.9%;	Score 1205;	DB 135;	Length 393;
Best Local Similarity	56.6%;	Pred. No. 0.00e+00;		
Matches 188;	Conservative	75;	Mismatches 67;	Indels 2;
			Gaps 2;	

Db 63 atgcaagagaagatgctcacttcatctatgttnccccaagaaggcg-gacctct 121  
||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :  
Qy 1 ATGCNNMGNRGARGAYCGNCAYTTTATHTAAGGNTAYCCNNAARAGCGNCAYGNCAYWSN 60

**Db**

122 tacaccacggctgaagaggccgctgggatcgcgccatcctcacagtcatcctggcgactcta 181  
||| || | ||| : ||| || | ||| : ||| : ||| : ||| :

**QV**

61 TAAACNAGCNGARGARCGNCNGGNA<sup>HGG</sup>NATHYTNACNGTNAVTHYNCCGTCTYTIN 120

Db  
182 cngctcatcggcgtattggtatnctagaaacgaatggatacagagcccttgatggataaa 241

Ov  
121 YTYNTNATGGCTGTCGTATYTGVMGMGMGNRAAYGCGTAYMGCNYTNATCGAAYAAR 180

Db 242 agtcttcattgttgcaactcaatgctgccttaacaagaagatgccacaagaagggtttgat 301  
:: : ||| || | ||| ||| : ||| : ||| ||| ||| ||| |||  
101 gsnvynicvctncnncanacadtccvcnvyniaannmcmntcyfncabarcgnntycav 240

[illegible]

Db 362 ccactgcttatgagaaactcttctgcagaac 393  
|| || || || || : || : || || :

## RESULT 6

DEFINITION	yx91a09.r1 Homo sapiens cDNA clone 269080 5' similar to PIR:A55253 melanoma antigen MART-1 - human ;.
ACCESSION	N36656

**KEYWORDS**  
EST.  
SOURCE  
human clone=26980 primer=T7 library=Soares melanocyte 2N8UM  
vector=pT7T3D (Pharmacia) with a modified polylinker host=DH101

cDNA was primed with a Not I - oligo(dT) primer  
[5'-TGTTACCAATCTGAAGTGGGAGCGCCGAGTTTTTTTTTTTTTTTT-3'],  
double-stranded cDNA was size selected, ligated to Eco RI adapted

RI sites of a modified pT7T3 vector (Pharmacia). Library (Pharmacia), digested with Not I and cloned into the Not I and

ORGANISM	Albino.
Homo sapiens	
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes	

REFERENCE

1 (bases 1 to 393)

Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Hominidae; Homo.

Sarcophagidae; Diptera; Sarcophaginae; Sarcophagini; Sarcophagus.

Ullrich, I. Clark N. Dubucque, T. Ellington K. Hawkins M.

Holman, M., Hultman, M., Kucaba, I., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and







Db 417 atccatcaaggctgtatccattgctctcataaccacacgagcggtgacag 473  
:||||| |: ||:|||| ||:|||| |: ||:||||:||||| |: ||: |: |  
Cp 177 RTCCATNAGCNGCKRTANCCRTTCKNCKNGKRCARTACCARCACCCDATNARNAR 121  
RESULT 10  
ID HS677260 standard; RNA; EST; 433 BP.  
AC N24677;  
DT 30-DEC-1995 (Rel. 46, Created)  
DT 30-DEC-1995 (Rel. 46, Last updated, Version 1)  
DE yx91a09.s1 Homo sapiens cDNA clone 269080 3'.  
KW EST.  
OS Homo sapiens (human)  
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;  
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.  
RN [1]  
RP 1-433  
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,  
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,  
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,  
RA Trevaskis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;  
RT "The WashU-Merck EST Project";  
RL Unpublished.  
CC Contact: Wilson RK WashU-Merck EST Project Washington University  
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,  
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:  
CC est@watson.wustl.edu High quality sequence stops: 315 Source: IMAGE  
CC Consortium, LNL This clone is available royalty-free through LNL  
CC ; contact the IMAGE Consortium (info@image.llnl.gov) for further  
CC information. NCBI gi: 1138827  
FH Key Location/Qualifiers  
FT 1..433  
FT source /organism="Homo sapiens"  
FT /clone="269080"  
FT /note="human"  
FT mRNA <1..>433  
SQ Sequence 433 BP; 119 A; 85 C; 86 G; 142 T; 1 other;  
Query Match 45.1%; Score 640; DB 135; Length 433;  
Best Local Similarity 56.8%; Pred. No. 1.17e-158;  
Matches 108; Conservative 36; Mismatches 45; Indels 1; Gaps 1;  
Db 244 ataagtggtgactgtcttcgcagagagttctcataagcaggtgagcattgggaac 303  
:||||| |: ||:|||| ||:|||| |: ||:||||:||||| |: ||: |: |  
Cp 348 RTANGGNGGNGWSYTGTCNGCNWSNARYTTTCTANGCNGGNGCCTTNGGNAC 289  
Db 304 cacaggttcacagttttctctgaagacacacttctgtcccgattgatcaaacccct 363  
|| ||:||||| |: ||:|||| |: ||:||||:||||| |: ||: |: |  
Cp 288 NACNGGYTCRCARTTTTTCYTCGNARNSNACYTTTWSRTCKCKRT-GRTCRAANCCYT 230  
Db 364 ctgtgggcatctcttctgtaagcacaattgagtgccaacatgaagacttttatccatca 423  
:||||| |: |: ||:|||| |: ||:||||:||||| |: ||: |: |  
Cp 229 CYTNGGRCACNCKNGKNGTARNRCRCAYTGNGTNGCNGACRTGNARNWSYTTTTCATNA 170  
Db 424 aggcctgtga 433  
: ||:|:  
Cp 169 RNCGNCKRTA 160  
RESULT 11  
LOCUS N24677 433 bp mRNA EST 28-DEC-1995  
DEFINITION yx91a09.s1 Homo sapiens cDNA clone 269080 3'.

ACCESSION N24677  
NID gi1138827  
KEYWORDS EST.  
SOURCE human clone-269080 primer=ml3 -40 forward library=Soares melanocyte  
2Nbhm vector=pf7T3D (Pharmacia) with a modified polylinker  
host=DH10B (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI Male.  
15'-TGTTACCAATCTAGCTGGAGCGCCGAGTTTTTTTTTTTTTTT-3'),  
double-stranded cDNA was size selected, ligated to Eco RI adapters  
(Pharmacia), digested with Not I and cloned into the Not I and Eco  
RI sites of a modified pf7T3 vector (Pharmacia). Library  
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal  
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.  
Albino.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;  
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;  
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 433)  
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,  
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,  
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and  
Wilson,R.  
TITLE The WashU-Merck EST Project  
JOURNAL Unpublished (1995)  
COMMENT Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
High quality sequence stops: 315  
Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
FEATURES  
source NCBI gi: 1138827  
Location/Qualifiers  
1..433  
/organism="Homo sapiens"  
/clone="269080"  
/note="human"  
mRNA <1..>433  
BASE COUNT 119 a 85 c 86 g 142 t 1 others  
ORIGIN  
Query Match 45.1%; Score 640; DB 113; Length 433;  
Best Local Similarity 56.8%; Pred. No. 1.17e-158;  
Matches 108; Conservative 36; Mismatches 45; Indels 1; Gaps 1;  
Db 244 ataagtggtgactgtcttcgcagagagttctcataagcaggtgagcattgggaac 303  
:||||| |: ||:|||| ||:|||| |: ||:||||:||||| |: ||: |: |  
Cp 348 RTANGGNGGNGWSYTGTCNGCNWSNARYTTTCTANGCNGGNGCCTTNGGNAC 289  
Db 304 cacaggttcacagttttctctgaagacacacttctgtcccgattgatcaaacccct 363  
|| ||:||||| |: ||:|||| |: ||:||||:||||| |: ||: |: |  
Cp 288 NACNGGYTCRCARTTTTTCYTCGNARNSNACYTTTWSRTCKCKRT-GRTCRAANCCYT 230  
Db 364 ctgtgggcatctcttctgtaagcacaattgagtgccaacatgaagacttttatccatca 423  
:||||| |: |: ||:|||| |: ||:||||:||||| |: ||: |: |





\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by Intelligenetics, Inc.

MPsrch nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Aug 27 08:43:36 1996; MasPar time 1419.91 Seconds  
1162.639 Million cell updates

Tabular output not generated.

Title: &gt;US-08-231-565A-26

Description: (1-2172) from US08231565A.seq

Perfect Score: 2172

N.A. Sequence: I GTCGAGGCCATACCAATC.....AAAAAAAAAAAAAAAAAAAA 21/2  
Comp: CAGCTGCCGGTAATGGTTAG.....TTTTTTTTTTTTTTTTTTTT

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 270440 seqs, 380027776 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: embl-new3

1:BCT 2:FUN 3:INV1 4:INV2 5:INV3 6:MAM 7:ORG 8:PLN  
9:PRI1 10:PRI2 11:PRO 12:ROD 13:SYN 14:UNC 15:VRT  
16:VTR1 17:VTR2

Database: genbank92

Database:

gerbank-new1

73:VRL4 74:VRL5 75:VRL6 76:VRL7 77:VRT1 78:VRT2 79:VRT3  
66:ROD7 67:STR 68:SYN 69:UNA 70:VRL1 71:VRL2 72:VRL3  
59:PRI13 60:ROD1 61:ROD2 62:ROD3 63:ROD4 64:ROD5 65:ROD6  
53:PR17 54:PR18 55:PR19 56:PR19 57:PR11 58:PR12  
46:PLN8 47:PR11 48:PR12 49:PR13 50:PR14 51:PR15 52:PR16  
39:PLN1 40:PLN2 41:PLN3 42:PLN4 43:PLN5 44:PLN6 45:PLN7  
32:INV6 33:MAM1 34:MAM2 35:PAT1 36:PAT2 37:PAT3 38:PHG  
25:BCR8 26:BCR7 27:INV1 28:INV2 29:INV3 30:INV4 31:INV5  
18:BCR1 19:BCR2 20:BCR3 21:BCR4 22:BCR5 23:BCR6 24:BCR7

Database: genbank-new1

87:ROD 88:STR 89:SYN 90:UNA 91:VRL 92:VRT  
80:BCT 81:INVI 82:INVZ 83:MAM 84:PHG 85:PLN 86:PRI

Database: u-emb145\_92\_11

y3:part1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

[illegible]

## ALIGNMENTS

KEYWORDS  
SOURCE  
ORGANISM

## SUMMARIES

Result No.	Score	Query		ID	Description	Pred. No.
		Match	Length			
1	2114	97.3	2130 59	ST3003	gpl00=melanocyte line	0.00e+00
2	2094	96.4	2114 55	HUMGPMS	Human 95 kd melanocyt	0.00e+00
3	1988	91.5	2026 52	HSGU1874	Human 95 kd melanocyt	0.00e+00
4	1974	90.9	2131 57	HUMPML	Human Pmel 17 mRNA, c	0.00e+00
5	1132	52.1	2817 86	HSPMEL17S2	Human melanocyte-spec	0.00e+00
6	1132	52.1	2817 10	HSPMEL17S2	Human melanocyte-spec	0.00e+00
7	806	37.1	1584 33	BOVRPELA	Bovine retinal pigmen	0.00e+00
8	752	34.6	1881 61	MMU14133	Mus musculus pmeil17 p	0.00e+00
9	726	33.4	2001 10	HSMEL30	Human melanocyte prot	0.00e+00
10	726	33.4	2001 86	HSPMEL03	Human melanocyte prot	0.00e+00
11	245	11.3	1178 10	HSMEL05	Human melanocyte prot	3.61e-110
12	245	11.3	1178 86	HSPMEL05	Human melanocyte prot	3.61e-110
13	209	9.6	1559 86	HSPMEL04	Human melanocyte prot	1.98e-90
14	209	9.6	1559 10	HSPMEL04	Human melanocyte prot	1.98e-90
15	150	6.9	673 86	HSPMEL02	Human melanocyte prot	1.46e-58
16	150	6.9	673 10	HSPMEL02	Human melanocyte prot	1.46e-58
17	88	4.1	577 10	HSPMEL01	Human melanocyte prot	2.58e-26
18	88	4.1	577 86	HSPMEL01	Human melanocyte prot	2.58e-26
19	88	4.1	3109 86	HSPMEL17S1	Human melanocyte-spec	2.58e-26
20	88	4.1	3109 10	HSPMEL17S1	Human melanocyte-spec	2.58e-26
21	46	2.1	1620 66	RNRNASTNV	R.norvegicus mRNA for 2	2.21e-06
22	45	2.1	1671 64	RATDBPA	Rat D-binding protein	5.98e-06
23	44	2.0	498 43	GLJAC01V	Glycine-rich protein	1.60e-05
24	44	2.0	987 40	GLJAC01V	Glycine-rich protein	1.60e-05
25	44	2.0	1710 39	BNPL2	B.napus (pNL2) mRNA f	1.60e-05
26	43	2.0	1934 79	S9519	stanniocalcin (oncorh	4.27e-05
27	43	2.0	2474 42	PEA70HSP	Plasium sativum chlorop	4.27e-05
28	43	2.0	2861 50	H58INDPR	H.sapiens mRNA for 2'	4.27e-05
29	43	2.0	3568 48	H525ABP	H.sapiens mRNA for 2-	4.27e-05
30	42	1.9	517 28	DDE428	Dictyostelium discoidl	1.13e-04
31	42	1.9	776 77	CHKTNC	Chicken tropinin C (T	1.13e-04
32	42	1.9	776 34	SSCKTIIB	Porcine CKII beta mRN	1.13e-04
33	42	1.9	837 59	S78214	APC=tumor suppressor,	1.13e-04
34	42	1.9	1042 59	S61953	c-erbB3= receptor tyro	1.13e-04
35	42	1.9	1305 61	MMYVD118	Mouse mRNA for Myd118	1.13e-04
36	42	1.9	1360 12	MM40808	Mus musculus glutathi	1.13e-04
37	42	1.9	1539 60	CGTUBB3	Gricetulus grieus (c	1.13e-04
38	42	1.9	1714 78	GUSDBF	C.domesticus mRNA for	1.13e-04
39	42	1.9	1777 30	EMCOLINF13	E.muelleri Lieb. COLN	1.13e-04
40	42	1.9	2062 45	WHTEF1X	Wheat translation elo	1.13e-04
41	42	1.9	3146 66	RRPCF3145	R.rattus mRNA for pot	1.13e-04
42	42	1.9	3426 81	DRDNPAT1	D.discoideum mRNA for	1.13e-04
43	42	1.9	3426 5	DRDNPAT1	D.discoideum mRNA for	1.13e-04
44	42	1.9	3442 12	RNRNAP81P	R.norvegicus mRNA for	1.13e-04
45	42	1.9	3504 29	DNKQBP1	Drosophila Knrl gene	1.13e-04

## ALIGNMENTS

## Unclassified.

REFERENCE 1 (bases 1 to 2130)  
AUTHORS Adema, G.J., de Boer, A.J., Vogel, A.M., Loonen, W.A. and Figdor, C.G.  
TITLE Molecular characterization of the melanocyte lineage-specific antigen gp100  
JOURNAL J. Biol. Chem. 269 (31), 20126-20133 (1994)  
REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gi: 639589] from the original journal article.  
COMMENT NCBI gi: 639589  
FEATURES  
    Location/Qualifiers  
        1..2130  
        /organism="Homo sapiens"  
        /note="human"  
        22..2007  
        /notes="melanocyte lineage-specific antigen/Pmel17 homolog; melanoma marker protein. Method: conceptual translation supplied by author. NCBI gi: 639590"  
        /codon\_start=1  
        /product="gp100"  
        /db\_xref="PID:g639590"

CDS  
    22..2007  
    /organism="Homo sapiens"  
    /note="human"  
    /notes="melanocyte lineage-specific antigen/Pmel17 homolog; melanoma marker protein. Method: conceptual translation supplied by author. NCBI gi: 639590"  
    /codon\_start=1  
    /product="gp100"  
    /db\_xref="PID:g639590"  
    /translation="MDLVKRLCLHLAVIGALLAVGATKVPKRNQDLGVSRQLRTKAWNRQLYEPWEAEQRLDCHRGQVSLKVSNDPPTLGNANFSALNFPSCQKVLPGQVIMVNTILNGSVWGQPVYQETIDDACIFPDGGPCPSGWSQKRSFYVYVWKTWQYVQLGCPVSLGISTGTRAMLTHTEVTVVHRRGSRVYPLAHSSNAFTITDQVPSVSVQLRALDGNKHFRLNQPLTFALQLHDPSCVLAENDLSYWDGDSSTGLISRALVYTHYILEPGVTAQVVAQAIPLTSCGSPSPVGTDDGHRPTAEAPNTTQAGVPTIEVGTTTQAPTAEPSTGTSVQVPTTEVTSTAPQMTAEISTGTEKVPVSEVNGTTLAEMSTPEATMPAEVSIIVLSCTTAQVTTTEWVETFARELIPEDGPDASSIMTSEITGSLGDLPGDATALRLKGVLPDLVLYRGSFVTLIDVQIESAEILQANPSGEDAFELTWSGGLPKEACMEISSPCQPAQLRCPVLPSPAQVLVHLKLGSGSTYCLNWSLADTNSLAVYSTOLLMPQOEAGLGVPLIVGILLVMAVLASLIYRRRLMKQDFSPVQLPHSSHMLRPLRFICSPIGENSPFLISGQV"

BASE COUNT 484 a 587 c 575 g 484 t

## ORIGIN

Query Match 97.3%; Score 2114; DB 59; Length 2130;  
Best Local Similarity 99.9%; Pred. No. 0.00e+00;  
Matches 2117; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 11 ggaagacacaatgatctggtgctaaaaagatgcctcttctatttggtgctgataggtg 70  
|||||  
Qy 28 GGAAGACACAATGATCGTGGTCTAAAGATGCGCTTCTTCAATTTGGCTGTGATAGGTG 87  
|||||  
Db 71 ctttgctggctgtgggggtacaaaagtacccagaaccaggactggttggtgtctcaa 130  
|||||  
Qy 88 CTTTGCTGGCTGTGGGGCTACAAAAGTACCAGAAACAGGACTGGCTGGTGTCTCAA 147  
|||||  
Db 131 ggaactcagaacaaagcctggaacaggcagctgtatccagagtggacagaagccaga 190  
|||||  
Qy 148 GGCACATCAGAACCAAGCCCTGCACAGGACGCTGTATCCAGAGTGGACAGAGCCAGA 207  
|||||  
Db 191 gacttgactgctggagagtggtcaagtgtccctcaaggtcagtaatgatggcctaac 250  
|||||  
Qy 208 GACTTGACTGCTGGAGAGGTGGTCAAGTGTCCCTCAGGTGAGTATGATGCGCCCTACAC 267  
|||||  
Db 251 tgattggtgcaaatgcctcctctctattgccttgaacttccctggagacaaaaggtat 310  
|||||  
Qy 268 TGAATGGTGCAAAATGCCTCCTCTATTTGCTTGAACCTCCCTGGAAAGCCAAAAGTAT 327  
|||||  
Db 311 tgccagatgggaggttatctggttcaacaataccatcatcaatgggagccaggtgtggg 370  
|||||  
Qy 328 TGCCAGATGGCAGGTATCTGGGTCAACAATACCATCATCAATGGGAGCCAGCGTGTGG 387  
|||||  
Db 371 gaggagagcagdgatgatccccaggaaactgacgatgcctgcatcttccctgatggtggac 430

Qy 388 GAGCAGCAGCAGTGTATCCCCAGAAACTGACGATGCGCTGCATTTCCCTCATGTGGTGCAC 447  
|||||  
Db 431 cttgccaatcgtggtctctcagaagaagagcttggttatgtctggaagacctggg 490  
|||||  
Qy 448 CTTGCCATCTGGCTCTTGGTCTCAGAAAGAAAGCTTTGTTATGCTCGAAGACCTGGG 507  
|||||  
Db 491 gccaatactgccaagtcttagggggccagctgtctgggtcagcattgggacagcaggg 550  
|||||  
Qy 508 GCCAATACTGGCAATTTCTAGGGGGCCACGTCTGCGCTGACGATTGGGACAGCAGGG 567  
|||||  
Db 551 caatgctgggcacacacaccatggaagtgaactgtctaccatgccggggatcccggagct 610  
|||||  
Qy 568 CAATGCTGGGCACACACACCATCGAAGTGACTGTCTTACCATGCCCGGGGATCCCGGAGT 627  
|||||  
Db 611 atgctccttctgctcattccagctcagccttcacattactgaccaggtgcctttctccg 670  
|||||  
Qy 628 ATGCGCTCTTGTCTCATTTCCAGCTCAGCTTACCATTTACTGACCAGGTGCTTCTCGG 687  
|||||  
Db 671 tgagcgtgtcccagttggyggccttggatggagggaacaagcactcctcgagaatcacgc 730  
|||||  
Qy 688 TGAGGCTGTCCAGTTGGGGCCCTTGGATGAGGGGACAAAGCACCTTCTCGAGAAATCAGC 747  
|||||  
Db 731 ctctgaccttggcctccagctccatgcccccagtggtctatctggctgaagctgacctct 790  
|||||  
Qy 748 CTCTGACCTTTGCCCTCCAGCTCCATGACCCGAGTGGCTATCTGGCTGAACTGACTCT 807  
|||||  
Db 791 cctacactggagacttgagacagtaagtgaacctgatctctcgggcacttgggtgta 850  
|||||  
Qy 808 CCTACACTGGGACTTTGGACACAGTACTGGAACCCCTGATCTCTCGGGCAGCTTGTGTCA 867  
|||||  
Db 851 ctcaatactactggagcctggccagctcactgccaggtggtcctgcaggtgcacattc 910  
|||||  
Qy 868 CTCATACTTACCTGGAGCTGGCCAGTCACTGCCAGGTGCTCTCGAGGCTGCCATTC 927  
|||||  
Db 911 ctctcacctcctgtggtcctccccagttccaggccacacagatgggcacaggccaactg 970  
|||||  
Qy 928 CTCTCACCTCTCTGGCTCTCCCGACTTCCAGGCACACAGATGGGCACAGGCCCACTG 987  
|||||  
Db 971 cagaggcccttaacaccacagctggccaaagtgcctactacagaagttggtgggtactaac 1030  
|||||  
Qy 988 CAGAGGCCCTTAACACCAACAGCTGGCCAAAGTGCCTACTACAGAAAGTTGTGGGTACTACAC 1047  
|||||  
Db 1031 ctggtcaggcgccaaactgcagagcctctggaaacacatctgtgcaggtgccaaacctg 1090  
|||||  
Qy 1048 CTGGTCAGGGCCCAACTGCAGAGCCCTCTGGAACACACATCTGTGCAGGTGCCAACCACTG 1107  
|||||  
Db 1091 aagtcataaagcactgcacctgtgcagatgccaactgcagagagcacaggtatgacacctg 1150  
|||||  
Qy 1108 AAGTCATAAGCACTGCACCTGTGCAGATGCCAACTGCAGACAGACAGATGACACCTG 1167  
|||||  
Db 1151 agaagtgccagtttcagaggtcatgggtaccacactggcagagatgtcaactccagag 1210  
|||||  
Qy 1168 AGAAGGTGCCAGTTTCAGAGGTGATGGGTACACACTGGCAGAGATGTCAACTTCCAGAGG 1227  
|||||  
Db 1211 ctacaggtatgacacctgcagaggtatcaattgtggtcttcttggaaacacagctgac 1270  
|||||  
Qy 1228 CTACAGGTATGACACCTGCAGAGGTATCAATTGCTGCTCTTCTTGAACCAACAGCTCAC 1287  
|||||  
Db 1271 aggtacaactacaggtgggtggagaccacagctagagagctacatcctcgagcctg 1330  
|||||  
Qy 1288 AGGTAAACAATACAGATGGGTGGAACCAACAGCTAGAGAGCTACCTATCCTTGCCTGAGCCTG 1347  
|||||  
Db 1331 aaggtccagatgccaggtcaatcatgtctacgggaagattatacaggttccctgggcccc 1390

Qy 1348 AAGGTCAGATGCCAGCTCAATCATGCTACGGAAAGTATTACAGGTTCCTCGGCCCC 1407  
 Db 1391 tgcgtgatgtacagccacacctaaagctggtagagacaagtcccccgtgattgttgc 1450  
 Qy 1408 TGCTGGATGGTACAGCACCTTAAGCTGGTGAAGACAGTCCCTCGATTTGTTTC 1467  
 Db 1451 tgtatcatatgtttcttccgtcacctggacatgtccaggtattgaaagtgcg 1510  
 Qy 1468 TGATCATATGTTCTTTTCGTACCTTGACATTTGTCAGGATTTGAAGTGGCG 1527  
 Db 1511 agatcctgcaggctgtgccctccgtgaggggatgcatattgagctgactgtgtcctgcc 1570  
 Qy 1528 AGATCCTGCAGGCTGTGCTGCGGTGAGGGGATGCATTTGAGCTGACTGTGTCTGCC 1587  
 Db 1571 aagcggtgtgccaaagagcctgcatggagatctcatcgcagggtggccagcccttg 1630  
 Qy 1588 AAGCGGGCTGCCAAGCAAGCTGATGGAGATCTCATCGCCAGGCTGCCAGCCCTG 1647  
 Db 1631 cccagcgtgtgccagctgtgtaccagccagcctgccagctgggtctgcaccaga 1690  
 Qy 1648 CCCAGCGCTGTGCCAGCTGTGCTACCCAGCCAGCTGCCAGCTGTGTCACACA 1707  
 Db 1691 tactgaaggtgtgctcgaggacatactgcctcaatgtgtctcgtgcatcaacagcc 1750  
 Qy 1708 TACTGAAGGTGGCTGGGGACATATGCTGCTCAATGTGCTGTGGTATACCAAGCC 1767  
 Db 1751 tggcagtggtcagaccagcttcatcagcctggtcgaagagagccttggcaggttc 1810  
 Qy 1768 TGGCAGTGTGCAGCACCCAGCTTATCATGCTGCTCAAGAGCAGGCTTGGCAGGTT 1827  
 Db 1811 cgcgtgctggtggcatcttgcgtgttgatggtgtggtccttgcatctctgatata 1870  
 Qy 1828 CGCTGATGCTGGGCATCTTGCTGCTGTGATGGCTGCTGCTGCTGATATATA 1887  
 Db 1871 ggcgcagacttatgaagcaagactctcctgacccagctgtccacatagcagctcact 1930  
 Qy 1888 GCGCAGACTTATGAAGCAGACTTCTCGTACCCAGCTGCCACATAGCAGAGTACT 1947  
 Db 1931 ggcgtcgtctacccgactcttgcctcgtcccatggtgagaatagcccccctcca 1990  
 Qy 1948 GGTGCTGTACCCGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2007  
 Db 1991 gt-gggcagcagcttgagctactctatgatgctgtgatttctcggagttgacagaaa 2050  
 Qy 2008 GTGGCAGCAGGCTGAGTACTCTCATATGATGCTGTGATTTCTTCCGTGGAGTTGACAGAAA 2067  
 Db 2051 cactctatttccccagcttctcctggagactactataactgaaataaataactcaga 2110  
 Qy 2068 CACCTATTTTCCCGCAGCTTCTCCCTGGGAGACTACTATTAACCTGAATAAATACTACA 2127  
 Db 2111 gctggaataaaaaa 2130  
 Qy 2128 GCTGAAAAAATAAAAAA 2147

RESULT 2  
 LOCUS HUMGPMSS 2114 bp mRNA PRI 15-JUN-1990  
 DEFINITION Human 95 kD melanocyte-specific secreted glycoprotein mRNA, 3' end.  
 ACCESSION M32295  
 NID g183559  
 KEYWORDS melanocyte-specific secreted glycoprotein.  
 SOURCE Human melanoma cell line, cDNA to mRNA, clone 8.  
 ORGANISM Homo sapiens

Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;  
 Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.  
 1 (bases 1 to 2114)  
 Vogel, A.  
 Sequence of a melanocyte specific secreted glycoprotein  
 Unpublished (1990) St. Louis Univ., St. Louis, MO 63104  
 Draft entry and computer-readable sequence for [1] kindly submitted  
 by A. Vogel, 23-FEB-1990.

NCBI gi: 183559  
 Location/Qualifiers  
 1..2114  
 /organism="Homo sapiens"  
 <1..2028  
 /note="melanocyte-specific secreted glycoprotein; NCBI  
 gi: 386754"  
 /codon\_start=1  
 /db\_xref="PID:g386754"  
 /translation="RGIRKNTMDIVLKRCLLHLAVIGALLAVGATKVPNRQDMLGYSR  
 QIRTKAMNRLYPTEAQRDLQMRGQVSLKVSNDGPTLIGANASFSIALNFPQSQK  
 VLPGQVIVWNTIINGSQWGGQVVPQETDDACIFPDGCPSCWSQKSFVYVW  
 KTWQYHQVLGCPYSGLSIGTGRMLGTHMTVYVHRRGSRVYLAHSSAFTLD  
 QVFSVSQIRALDGGKHFLRNQPLTFALQLDHPSGYLAEDLSYTWDFGDSGTL  
 ISRALVYTHYLEP GPVTAQVLIQAAIP LITSCSSSPVPTDGHRTAEAPNTTAGOV  
 PTEVVGTPGQAPTAEPSCTTSVQPTTEVISTAPVQMPTEASTMTPKEKVPSEVM  
 GTTLAEMSTPEATGMTAEVSVIVLSTAAQVTTTETWETARELP IPEPEGRDASS  
 IMSTESITGSLGLDLDGTATLAKRVQPLDCVLYRYSFVTLDIVQGESAEILQA  
 VPSGEGDAFELTVSCGELPKACMEISSPGQPPAORLQVLPSPACOLIVLHQLK  
 GSGTGYCIANSLADNSLAVSTQLIMPQQEAGLQVPLVIGLILVLAHVLAFLIYR  
 RRLMKQDFVPLP LPHSSSHLRLPASALLPLVIRAFSSVSRSEYSHMML"

BASE COUNT 469 a 586 c 575 g 484 t  
 ORIGIN

Query Match 96.4%; Score 2094; DB 55; Length 2114;  
 Best Local Similarity 99.9%; Pred. No. 0.00e+00;  
 Matches 2102; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
 Db 11 ggaagacacacatgagctgggtctaaaaagatgcctctcttcatttgctgtagatggg 70  
 Qy 28 GGAAGAACACATGGATCTGCTGCTAAAAAGATGCCCTCTTCATTTGGCTGATAGTG 87  
 Db 71 ctttgctgctgtgggggtacaaagtaccagaaaccagagactggtggtgtctcaa 130  
 Qy 88 CTTTGCTGGCTGTGGGGCTACAAAAGTACCAGAAACCCAGACTGGCTTGGTCTCTCA 147  
 Db 131 ggaactcagaacaaagcctggacagcagctgtatccagagtggacagaagccaga 190  
 Qy 148 GCAACTCAGAACAAAGCCTGGAAAGCAGGAGCTGTATCCAGAGTGAGACAGACCCACA 207  
 Db 191 gacttgactgctggagagtggtccaagtgtccctcaagtcagtaaatgatggcctaac 250  
 Qy 208 GACTTGACTGCTGGAGGTGGTCAAGTGTCCCTCAAAGTCAAGTCAATGATGGGCTACAC 267  
 Db 251 tgattggtgcaaatgcctcctctctcttgccttgacttccctggaagccaaaggtat 310  
 Qy 268 TGATTGGTGAATGCCTCTCTCTATTGGCTTGAATCTCCCTGGAGCCAAAGGTAT 327  
 Db 311 tgcagatggcaggttatctgggtcaacaataaccatcatcaatgggagccaggtggg 370  
 Qy 328 TGCAGATGGCAGAGTTATCTGGGTCAACAATACCATCATCAATGGGAGCCAGGTGTGG 387  
 Db 371 gaggacagcaggtgtatccccaggaactgacgatgcctgcctcctcctgatgggtggac 430  
 Qy 388 GAGGACACCCAGTGTATCCCGACGAAACTGACGATGCTTCCCTCATGCTGCTGCTG 447







REFERENCE 1 (bases 1 to 2023)  
AUTHORS Maresch, G.A., Marken, J.S., Neubaumer, M., Aruffo, A., Hellstrom, I.,  
Hellstrom, K. and Marquardt, H.  
TITLE Cloning and expression of the gene for the Melanoma-Associated ME20  
Antigen  
JOURNAL DNA Cell Biol. 13, 87-95 (1994)  
MEDLINE 94235165  
REFERENCE 2 (bases 1 to 2026)  
AUTHORS Neubaumer, M.G.  
TITLE Direct Submission  
JOURNAL Submitted (16-SEP-1993) Michael G. Neubaumer, Bristol-Myers Squibb  
Pharmaceutical Research Institute, 3005 1st Ave, Seattle, WA 98121,  
USA

COMMENT NCBI gi: 494939

FEATURES Location/Qualifiers  
source

1..2026  
/clone="hf12-2"  
/clone\_lib="cdm8"  
/organism="Homo sapiens"  
/cell\_line="melanoma H3606"  
7..78  
7..1995  
/standard\_names="melanoma-associated ME20 antigen"  
/note="NCBI gi: 494940"  
/codon\_start=1  
/product="me20m"  
/db\_xref="PID:g494940"

sig\_peptide

CDS

/translation="MDLVKRCILHLAVIGALLAVCAIKVPRNQDLWGLVSRQLRTKAW  
NHQLYPEWTEAQRIDCRGQVSLKVSNDGPTLIAGNASFSLAINPFGSKVLPDQV  
IWNNTIINGSOVWGPVYPOETDDACIFDGPQPCSGWSOKRSFYVWKTWGQYW  
QVLCGPVSGLSIGTCRAMLGTHMEVTVYHRRGSRSPYLAHSSNAFTIDQVPSVS  
VSQLRDLGGNKHFIRNQPLTFALQHDPSGYLAELADLSYTWDFGDSGSLISRALV  
THYLEPQPVTAQVVLQALIPITSCGSPVPVTTDCHRTAEAPNTAGQVPTTVVG  
TTPQATPAEPSCITSVQVTEVISTAPVQMTAEFTGTPKPYVSEVWGTTLAEM  
STPEATCMTPAESIVLISGTAQVTTTETVETARELP IPEPEGPASSIMSTESI  
TGSGLPLDGTATLRVLRQVPLDCLVLYRGFSVTLIDVQIESAEILQAPVSEGD  
AFELTVSCGGLPKPEACMEISSPGQPPAQLCPVLPSPACOLVHOLIKGSGTYC  
LNVSLADNLSLAVSTOLIMPQOEAGGLGQVPLIVGILLVLMVAVLSLIYRRLMKQ  
DFSPVQLPIHSSHLRLPRIFCSCPIGENSPLISQQV"

mat\_peptide

/product="me20m"  
BASE COUNT 437 a 564 c 564 g 461 t  
ORIGIN

Query Match 91.5%; Score 1988; DB 52; Length 2026;  
Best Local Similarity 99.8%; Pred. No. 0.00e+00;  
Matches 2008; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

Db 7 atggatcgtgggtctaaaagatgcctcttcatttgctgtgataggtgtcttgctggct 66  
|||||  
Qy 39 ATGGATCTGGTGGTAAAAGATGCCTCTTCATTTGGCTGTGATAGTCTTGTGGCT 98  
|||||  
Db 67 gtgggggtcacaaaagtaccagaaccaggactgcttggctgtctcaaggcaactcaga 126  
|||||  
Qy 99 GTGGGGGTACAAAAGTACCAGAAACCAGGACTTGGTGTCTCAAGGCAACTCAGA 158  
|||||  
Db 127 accaaagcttgaaacaggcagctgtatccagagtggacagaagcccgagagactgaactgc 186  
|||||  
Qy 159 ACCAAAGCTGGAACAGGCAGCTGTATCCAGGTGGACAGAAAGCCCGACAGACTTGA 218  
|||||  
Db 187 tggagaggtgggtcaagtgtccctcaagtcagtaagtatgggctacactgattgtgca 246  
|||||  
Qy 219 TGGAGAGGTGCTCAAGTCTCAAGTCAAGTAATGATGGGCTACACTGATTGGTGA 278  
|||||

Db 247 aatgcctcctctctctatgtccttgaacttccctggaagccaaaggtattgcccagatggg 306  
|||||  
Qy 279 AATGCCCTCTCTCTATTGCCCTTCAACTTCCTGGAGCCAAAAGGTATTGCCAGATGGG 338  
|||||  
Db 307 caggttatctgggtcaacaataccatcatcaatggagccagggtgtgggagggacagcca 366  
|||||  
Qy 339 CAGGTATTCTGGGTCAACANATACCATCATCAATGGAGCCAGGTGTGGGAGGACAGCCA 398  
|||||  
Db 367 gttatccccaggaaaactgacgatgcctgcctcttccctgatggtgaccttgccctatct 426  
|||||  
Qy 399 GTGTATCCCCAGGAAACTGACGATGCCTGCATCTCCCTGATGTGGACCTTGGCCCATCT 458  
|||||  
Db 427 ggctcttggtctcagaagaagactttgtttatgtctggaagacctggggccaataactgg 486  
|||||  
Qy 459 GGCTCTTGCTCTCAGAAGAGAAAGCTTTGTTTATGTCTGGAAGACTGGGGCCAACTACTGG 518  
|||||  
Db 487 caagttctagggggcccaagtgtctgggtgagcaattgggacagggcagggcaatgctgggc 546  
|||||  
Qy 519 CAATTTCTAGGGGGCCCACTGTCTGGGCTGAGCATTTGGGACAGCGAGGCCAATGCTGGGC 578  
|||||  
Db 547 acacacacatggaagtgaactgtctaccatcgccggggtatcccgaggtatgtgctcttt 606  
|||||  
Qy 579 ACACACACATGGAGACTGACTGTCTACCATCGCGGGGATCCGGAGCTATGTGCCTCTT 638  
|||||  
Db 607 gctcattccagctcagccttcaccattactgaccaggtgctcttctccgtgagcgtgtcc 666  
|||||  
Qy 639 GCTCATTCAGCTCAGCTTCACCATTAATGACCCAGGTGCTTCTCCGTGAGCGGTGCC 698  
|||||  
Db 667 cagttcggggccttggatggagggaagaacacttctctgaaataacagcctctgacctt 726  
|||||  
Qy 699 CAGTTCCGGGCTTGGATGGAGGGAACANGACACTTCTCGAATAATCAGGCTCTGACCTTT 758  
|||||  
Db 727 qcctccagctccatgacccagtggtctatctggctgaagctgacctctctacacactgg 786  
|||||  
Qy 759 GCCTCCAGCTTCATACCCAGTGGCTATCTGGCTGAAGCTGACCTCTCCACACCTGG 818  
|||||  
Db 787 gacttggagacagtgtggaacctgatctctcgggcaacttgggtcaactactaac 846  
|||||  
Qy 819 GACTTTGGACAGCAGTAGTGAACCCCTGATCTCTCGGGCACTTGTGGTCACTCATCTTAC 878  
|||||  
Db 847 ctggagcctggggcagctcaactgcccaggtggctgagcaggtgcccattcctcctcacctcc 906  
|||||  
Qy 879 CTGGAGCCTGGCCCACTCACTGCCAGGTGGTCTCTGAGGCTGCCATTCTCTCACTCC 938  
|||||  
Db 907 tgtggctctccccagttccaggccacacagatgggacaggcccaactgcagagggccct 966  
|||||  
Qy 939 TGTGGCTCTCTCCCACTTCAGGGCACACAGTGGGCACAGGCCCACTCGACAGGCCCT 998  
|||||  
Db 967 aacacacacagctggccaaagtgcctactacagaagttgtgggtactacacctgggtcagggc 1026  
|||||  
Qy 999 AACAACACAGCTGGCCAAAGTGCCTACTACAGAGTTGTGGGTACTACACTGGTCAGGGC 1058  
|||||  
Db 1027 ccaactcgagagccctctggaaccacatctgtgaggtgccaaccactgaagtcataaagc 1086  
|||||  
Qy 1059 CCAACTGGACAGGCCCTCTGGAACCACTCTCTGAGGTGCCAACCACTGAAGTCAATAGC 1118  
|||||  
Db 1087 actgacacctgtgaagatgccaactgcagagagcacaggtatgacacctgagaaggtgcca 1146  
|||||  
Qy 1119 ACTGCACCTGTGCAGATGCCAACTGCACAGAGCACAGGTATGCACACTCAGAGGTGCCA 1178  
|||||  
Db 1147 gtttcagaggtgcatgggtaccacactggcagagatgtcaactccagaggtacaggtatg 1206  
|||||  
Qy 1179 GTTTCAGAGGTCAATGGGTACCACTGGCCAGATGCACTCCAGAGGCTACAGGTATG 1238  
|||||

Db	1207	acacctcgadaggtatcaatttgytgcttcttgaaaccaacagctgcacagtgtaaacact	1261
Qy	1239	ACACCTGCAGAGGTATCAATTGTGGTGGCTTTCTGGAAACACAGCTGCACAGTTAACAACT	1298
Db	1267	acagagtgggtggagaccacagctagagagctacctatccctgagcctgaaggtccagat	1326
Qy	1299	ACAGAGTGGGTGGA GACCACAGCTAGACAGCTACTTACCTCCCTGAGCCTGAGAGTCCAGAT	1358
Db	1327	gccagctcaatcatgtctacgaaagattacaggttccctgggcccctctgctggatggt	1386
Qy	1359	GCCAGCTCAATCATGTCTACGGAAAGTATTACAGGTTCCCTTGGCCCCCTGCTGGATGGT	1418
Db	1387	acagccacctaaagctgtggaagagacaagtcctccctggatgtgtctctatccgatat	1446
Qy	1419	ACAGCCACTTAAGCTGGTGAAGACAGACATCCCTCGGATTGTCTTGTATTCGATAT	1478
Db	1447	ggttcctttccgtcacccctggacattgtccaggttatgaaagtcgcagagatccctgacg	1506
Qy	1479	GGTTCTCTTTCCGTCA CCTGGACATCTCTCAGGGTATTGAAGTGCAGATCCTCGAC	1538
Db	1507	gctgtccctccggtgagggggatgcatctgcagaggtgccagccccctgccacaggcggctg	1566
Qy	1539	GCTGTGCCGTCCGGTCAGGGGGATGCATTTTCAGCTGACTGTGTCTCTGCCAAGGGGGCTG	1598
Db	1567	cccaaggaagcctgcattggagatctcatgccagaggtgccagccccctgccacagcggctg	1626
Qy	1599	CCCAAGGAAGCTGCATGCGAGATCTATCGCGAGGGTGCACAGCCCCCTGCCACGGGCTG	1658
Db	1627	tgccagcctgtgtaccacagccagcgtccagcgtggtctgcaccagatactgaaggt	1686
Qy	1659	TGCGAGCCTGTGTACCCAGCGCAGCGCTGCCAGCTGCTTCTGCACAGATACTGAGGGT	1718
Db	1687	ggctcggggacatacgtccataatgtctctaggtgtatcccaacagcctgcagtggtc	1746
Qy	1719	GGCTCGGGGACATATCGCTCAATGTGTCTCTGGCTGATCCAAACAGCCTGGCAGTGGTCT	1778
Db	1747	agcaccacagcttatcatgcctggtcaagaacagggggccttgggcaggttcgcgtgac	1806
Qy	1779	AGCACCCAGTTATCATGCTTGCTCAAGACGAGG----CCTTGGGACAGGTTCCGCTCATC	1835
Db	1807	gtggacatctgctgggtgttgatggctgggtccttgcatctctgatataagggcaga	1866
Qy	1836	GTGGGATCTTGTGCTTTGATGGCTGTGGCTCTTTGCATCTCTGATATATAGCGCAGA	1895
Db	1867	cttatgaagcaagaactctccgtaccacaggtgccacatagcagcagtcactggctggct	1926
Qy	1896	CTTATGAAGAGACATTCCTCTGGTATCCCAAGTTCGACATACGACAGTCACTGGCTGGCT	1955
Db	1927	ctaccgccatctctgctctgtgccattggtagaataagccccctctcagtgggcag	1986
Qy	1956	CTACCCCGGCATCTTCTGCTCTTCTCCCAATTGGTGAGACAGCCCCCTCTCAGTGGGCAG	2015
Db	1987	caggtctgaatcactctcatatgactgtgatt	2019
Qy	2016	CAGGTTCTGACTCTCTATATGCTGTGATT	2048

RESULT	4	08-JAN-1995
LOCUS	HUMPMEL	2131 bp mRNA
DEFINITION	Human Pmel 17 mRNA, complete cds.	
ACCESSION	M77348	
NID	g190105	
KEYWORDS	Pmel 17 protein; melanocyte.	
SOURCE	Homo sapiens skin cDNA to mRNA.	

ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea; Hominidae.
AUTHORS	1 (bases 1 to 2131) Kwon,B.S., Chintamaneni,C., Kozak,C.A., Copeland,N.G., Gilbert,D.J., Jenkins,N., Barton,D., Francke,U., Kobayashi,Y. and Kim,K.K.
TITLE	A melanocyte-specific gene, Pmel 17, maps near the silver coat color locus on mouse chromosome 10 and is in a syntenic region on human chromosome 12
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 88 (20), 9228-9232 (1991)
MEDLINE	92021023
COMMENT	NCBI gi: 190105
FEATURES	Location/Qualifiers 1..2131 /organism="Homo sapiens" /cell_line="primary culture" /cell_type="melanocyte" /sequenced_mol="cDNA to mRNA" /tissue_type="skin" 1..2121 /gene="Pmel 17" 12..2018 /gene="Pmel 17" /note="NCBI gi: 190106" /codon_start=1 /product="Pmel 17 protein" /db_xref="PID:g190106" /translation="MDILVLRKCLLHLAVIGALLAVGATKVPKNODMLGVSRQLRTKAW NRQLPWTAEQRDLDRNRGGQVSLKNSNDGPTLGNANFSIALNFFGQSKVLPDQVW INWNTTINQSRQWGGQVLPQETDADCFDQGPCEPGSMQSKRSEVYKMTQWQV QVLGGPVGSLSTCTGRAMLCTHMTVTVYRRGSRVYPLAHSSAFPTTDQVPEFS VSQRLADGNKHFRLNQLPTFLQLHDPSSCYLAEDLSYTWDFGSSGCTLIISRAPV THYLEPQVPTAVLQQAIPITSCGSSPVTCTDGHRTAEAPNTAGQVPTTEVVG TTPQACGPEAGTTSVTSVTTAEVITAPVQMPTEASTGMTPEKVPSEVMGTTIAEM SPTEATCWTPAEVSIIVLGSTTAQVNTTETWETTARELP IPEKGPDASSIMTESI TCSLQPLDGTATRLVKQVPLDCVLVRYGSFVTLIDVQGIESAELIQAVPSGED APFLVTSGLGPEAKCAEMISSPQCPAPRIQCQVILPSPACQLVHLHQILKGGSGTYC LNVLSDATNSLAVYSTQLMPVGLILITQEAGLQGVRLVIGILLVLMVAWLASLIAYR RELKMQDFSVPLQPHSSSHLRLPRIFCSCPIGENSPILSGQV"
BASE COUNT	475 a 588 c 578 g 490 t
ORIGIN	

Query Match	90.9%;	Score 1974;	DB 57;	Length 2131;
Best Local Similarity	98.8%;	Pred. No. 0.00e+00;		
Matches 2105;	Conservative	0;	Mismatches 5;	Indels 21;
				Gaps 1;
Db	1	ggaagaacaacatgctgctgctaaaaagatgcctcttcatttgcctgtgataggtg	60	
Qy	28	GGAAAGAACAAATGGATCTGTGCTTAAAGATGCGCTTCTCATTTGGCTGTCATAGGTG	87	
Db	61	ctttgcctgcctgctgggggctacaaaagtaccagaaaccaggactgctggtgtctcaa	120	
Qy	88	CTTTGCTGGCTGTGGGGGCTACAAAGTACCCAGAAACAGGACTGGCTGGTGTCTCAA	147	
Db	121	ggcaactcagacaacaaagctggaaacaggcagctgtatccagagtggaacagaaagccacga	180	
Qy	148	GGCAACTCAGAAACCAAAAGCTCGAAACAGGCAGCTGTATCCAGAGTGGACAGAAGCCACGA	207	
Db	181	gacttgactgctggagaggtggctcaagtgccctcaagtgtcagtaatgatgggctctaac	240	
Qy	208	GACTTGACTGCTGGAGAGGTGGTCAAGTGTCCCTCAGGTCAGTAATCATGGCGCTACAC	267	
Db	241	tgatttgagcaaatgcctctctctctattgccttgaacctccctggaaagccaaaagttat	300	

QY 268 TGATTGGTCAAAATGCCTCTCTATTGGCTTGAACTTCCCTGGAAGCAAAAGGTAT 327  
DB 301 tgccagatggcgaggttatctgggtcaacaataccatcatcaatggagccaggtgtggg 360  
QY 328 TCCAGATGGGAGGTTATCTGGGTCAACATACCATCATCAATGGGAGCCAGGTGGG 387  
DB 361 gaggacagccagtgatcccccaggaaactgaagatgctctccctgagtggtggac 420  
QY 388 GAGGACAGCCAGTGATGCCCGAGAAATCAGGATGCCCTGCATCTTCCCTGATGGTGGAC 447  
DB 421 ctggccatctggctcttggtctcaagaagaagctttgttatgtctggaagacctggg 480  
QY 448 CTGGCCATCTGGCTCTTGCTCTCAGAGAGAAAGCTTTGTTATGCTCGAAGACCTGGG 507  
DB 481 gccaatactggcaagtctagggggcccagtgctgggctgagcattgggagcaggcaggg 540  
QY 508 GCCAATACTGGCAATTTCTAGGGGGCCAGTGTCTGGGCTGAGCATTTGGGACAGGACGG 567  
DB 541 caatgctgggacacacacacatggaagtgactgtctaccatgccggggggtatcccggagct 600  
QY 568 CAATGCTGGGACACACACATGGAAGTGACTGTCTACCATGCCGGGGATCCGGAGCT 627  
DB 601 atgtgctctgtgctcattccagctcagccttcaccattactgaccagtgctcttctccg 660  
QY 628 ATGTGGCTCTTGCTCATTCAGCTCAGCCTTCACCATTAAGGAGGTGCTTCTCCG 687  
DB 661 tgagcgtgtcccaggttccaggttccaggtccaggtgagaggaacaagaacttctgagaaatcagc 720  
QY 688 TGAAGCTGTCCCAGTTGCGGGCTTGGATGGAGGGAACAAGCACTTCTCGAANAATCAGC 747  
DB 721 ctctgaccttggcctccaggtccatgacctagtgtctatctggtgagtgagctgacctct 780  
QY 748 CTCTGACCTTTGGCTTCAGCTCCATGACCCAGTGGCTATCTGTGCTGACGTGACCTCT 807  
DB 781 cctaaccttgggaacttggagacagtagtggaacctgtatctctcgggcaactgtggtca 840  
QY 808 CTTACACTGGGACTTTGGAGACACTAGTGGAAACCTGTATCTCTGGGCACTTGTGGTCA 867  
DB 841 ctcatacttacctgagcctggccagtcactgcccaggtggtcctcgaggtgcccattc 900  
QY 868 CTCATACTTACCTGGAGCCTGGCCCACTGCTGCCAGGTGGCTGCGAGGCTGCCATTC 927  
DB 901 ctctcaactctgtgctcctcccagttccaggccacacagatgggcaagcccaactg 960  
QY 928 CTCTCACTCTGTGGCTCTCCCATGTTCCAGGCAACAGATGGGACAGGCCAACTG 987  
DB 961 cagaggccctaacacacagctggccaagtgcctactacagaagtgtggtgactaac 1020  
QY 988 CAGAGGCCCCCTAACACACAGCTGGCCAACTGGCTACTACAGAAGTTGTGGGTACTAC 1047  
DB 1021 ctgggtcagggcccaactgcagagccctctggaaccacatctgtgaggtgccaaacctg 1080  
QY 1048 CTGGTCAAGGGCCCACTGCAGAGCCCTCTGGAAACCATCTGTGCAAGTGGCAACCACTG 1107  
DB 1081 aagtcataagcactgacctgtgagatgccaaactggaagagacaggtatgaaacctg 1140  
QY 1108 AAGTCAATAGCACTGCACCTGTGCAATGGCAATCTGCAGAGACAGGATATGACACCTG 1167  
DB 1141 aagaaggtgccaggttccagaggtcatgggtaccacactggcagagatgtcaactccagagg 1200  
QY 1168 AGAAGGTGCCAGTTTCAGAGGCTCATGGCTACCACATGGCAGAGATGTCAACTCCAGAGG 1227  
DB 1201 ctacaaggtatgacacctgcagaggtatcaattgtgggtcttcttggaaaccacagctgcac 1260

QY 1228 CTACAGGTATGACACTGCAGAGGTTATCAATTGTGTGCTTTCTGGAACACAGCTGCAC 1287  
DB 1261 aggttaacaactacagagtgggtggagaccacagctagagagtaacctatccctggagcctg 1320  
QY 1288 AGGTAAACAATACAGACTGGGTGGAGACACAGCTAGAGAGCTTACCTATCCCTGAGCCTG 1347  
DB 1321 aaggtccagatgccagctcaatctctacggaaagtattacaggttccctggggccccc 1380  
QY 1348 AAGTTCAGATGCCAGCTCAATCTGTACGGAAGTATTACAGGTTCCCTGGGCCCCC 1407  
DB 1381 tctctggatggtacagccacacttaagctgggtgaagacacaagtccccctggattgttgc 1440  
QY 1408 TGCTGATGTGTACAGCCACCTTAAGGCTGGTGAAGACAAGTCCCCCTGGATTTGTGTTTC 1467  
DB 1441 tgtatcatatggttcttccgtccacctggacattgtccaggggtattgaaagtgcgc 1500  
QY 1468 TGTATCGATATGTTCCCTTTTCCGTCACTCCCTGGACATTTGTCCAGGCTATTGAAAGTGGCG 1527  
DB 1501 agatcctgcaggtgtgcggtccggtgagggggatgcatttgagctgactgtctcctgcc 1560  
QY 1528 AGATCTGCAGGCTGTGCCGTCCGGTGAGGGGGATGCATTTGAGCTGACTGTCTCTGCC 1587  
DB 1561 aaggcgggctgcccgaagagcctgcattgagatctcatgccaggggtgccagcccccctg 1620  
QY 1588 AAGCGGGGCTGCCCAAGGAAGCTGCATGGAGATCTCATGCCAGGGTGCACGCCCTCG 1647  
DB 1621 cccagcggctgtgccagcctgtgctaccagcccagcctgccagctggttctgcaccaga 1680  
QY 1648 CCCAGCGGCTGTGCCAGCCTGTGCTTACCAGCCAGCCTGCCAGCTGGTTCGCAACGAGA 1707  
DB 1681 tactgaaggggtgggtcggggacatactgctcaatgtgctctggctgataaccaacagcc 1740  
QY 1708 TACTCAAGGTTGGCTCGGGGACATACTGCCTCAATGTCTCTGTGGCTGATACCAACAGCC 1767  
DB 1741 tggcagtggtcagcagccagcttatcagcctgtcctgggattcttctcacaggtcaag 1800  
QY 1768 TGGCAGTGGTGCAGACCCAGCTTATCATGCTG-----GTCAAG 1806  
DB 1801 aaqcaagccttggcgaggttcgggtgatcgtgggcatcttgggtgtgatgggtgtgg 1860  
QY 1807 AAGCAGGCTTGGGCGAGGTTCCGCTGATTCGTGGGCATCTTGTGGTGTGATGGCTGTGG 1866  
DB 1861 tcccttgcatctctgatataataggcgagacttatgaagcaagacttctccgtaccaccagt 1920  
QY 1867 TCCTTGCATCTCTGATATATAGGCGCAGACTTATGAGCAAGACTTCTCCGTACCOCAGT 1926  
DB 1921 tgccacatagcagcagtcactgggtgogtctaccocgactcttctgtcttgtccocattg 1980  
QY 1927 TGCCCATATGACAGCTCAGTGGCTGGGTCTTACCCCGCATCTTCTGCTCTTGTGCCATTG 1986  
DB 1981 gtgaagaatagccccctcctcagtgggcagcaggtctgagtaactctcatatgatgtgtga 2040  
QY 1987 GTGAGACACAGCCCCCTCTCTCAGTGGGAGCAGGCTGTGAGTACTCTCTCATATGATGCTGTGA 2046  
DB 2041 ttttctcggagttgacagaacaacctatatttccccagctcttccctggagagactactat 2100  
QY 2047 TTTTCTCGGAGTTGACAGAAACACCTATATTTTCCCCCAGTCTTTCCCTGGGAGACTACTAT 2106  
DB 2101 taactgaataaataactcagagcctgaaaaa 2131  
QY 2107 TAACGTAAATAAATACTCAGAGCCTGAAGAAA 2137



Db 2105 cagggtattgaagtgccagagatcctgcaggctgtgccgtccggtgaggggatcattt 2164  
|||||  
Qy 1509 CAGGGATTGAAGTCCCGAGATCTCGAGGCTGTGCCGTCCGCTGAGGGGATCATTT 1568  
|||||  
Db 2165 gacgtgaetgtgtcctgccaaaggcgtgccaaaggaagcctgcagtgagatctcatcg 2224  
|||||  
Qy 1569 GAGCTGACTGTGTCTGCCAAGGGGGCTGCCAAGAGAGCCTGCTGAGAGATCTCATCG 1628  
|||||  
Db 2225 ccagggtgccagccctcccaagcgtgtgccagcctgtgctaccagccagcctgc 2284  
|||||  
Qy 1629 CAGGGTGCAGCCCTCCAGCGGCTGTCCAGCCTGTCTACACCCAGCCTGC 1688  
|||||  
Db 2285 cagctggttgacagagatactgaaggtggctcggggacatactcctcaatgtgtct 2344  
|||||  
Qy 1689 CAGCTGCTTCCACAGACTACTGAAGGCTGCTCGGGGACATACCTCCTCAATCTGTCT 1748  
|||||  
Db 2345 ctggtgataccacagcctggcagtggtcagcaccagccttatcatgcctggt 2398  
|||||  
Qy 1749 CTGCTGATACACAGCCTGGAGTGTTCAGCACCACCTTATCATGCTGCT 1802  
|||||

## RESULT. 6

ID HSMELI752 standard; DNA; PRI; 2817 BP.  
AC U20093;  
DT 01-JAN-1996 (Rel. 46, Created)  
DT 01-JAN-1996 (Rel. 46, Last updated, Version 1)  
DE Human melanocyte-specific (pmel 17) gene, exons 2-5, and complete cds.

OS Homo sapiens (human)  
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;  
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.

RN [1]  
RP 1-2817  
RA Kwon B.S., Kim K., Heng H.H., Shi X.M., Teui L., Lee Z.H.,  
RA Youn B., Pickard R.T.;  
RT "Genomic organization and FISH mapping of human pmel 17, the  
RT putative silver locus";  
RL Unpublished.

RN [2]  
RP 1-2817  
RA Kwon B.S.;  
RT ;  
RL Submitted (05-JAN-1995) to the EMBL/GenBank/DBJ databases.  
RL Byoung S. Kwon, Indiana University School of Medicine, Microbiology  
RL and Immunology, 635 Barnhill Drive, Indianapolis, IN 46202, USA  
CC NCBI gi: 1142634

FN Key Location/Qualifiers

FT source 1..2817  
FT /clone\_lib="lambda FIX II from Stratagene"  
FT /organism="Homo sapiens"  
FT /chromosome="12"  
FT /map="12pter-q21"  
FT join(019491:1111..1186,67..459,853..1014,1266..2396,2499..  
FT 2722)  
FT /gene="pmel 17"  
FT /codon\_start=1  
FT /product="Pmel 17"  
FT /db\_xref="PDB:gi1142636"  
FT /translation="MDLVKRLCLHLAVIGALLAVGATKVPKPNQDMIGVSRQLATKAW  
FT N  
FT RQLYFWEFAQRDLCHRGGQVSLKVSNDGPTLIGANASFTALNFPCKVLPDQVIT  
FT W

FT VNNIINGSQVWGQVVPQETDDACIFPDGCPSCSHSQKRSFYVYWKTWQVWQV  
L  
FT GGPVSGLSITGTGRAMLGHTMEVTVVHRRGSRSVPLAHSSSAFTITDQVPFVSVSQ  
L  
FT RALDGNKHFRLNQPLITFALQLHDPSCYLAEDALSYTWNDFGDSSGTLISRAPVWTHY  
L  
FT EPGVTAQVVLQAAIPLISCGSPVPTDGHRTAEAPNTTAGQVPTTEVVGTPCQ  
A  
FT PTAEPSCGTTSVQVPTTEVISTAPVQMPAEATGCTPEKVPVSEVMGTTLAEMSTPEAT  
G  
FT MTPAEVSVILSGTTAAQVTTTEWETTARELP IPEPEGPDASSIMSTESITGSLGPL  
L  
FT DGTATLRLVRQVPLDCLVLYRYSFSVTLDIVQGISAEILQAVPSGEGDAFELLTVC  
Q  
FT GGLPEACMEISSPGQPAPQLRQCPVLPSPAQLVHLQILKGGSGTYCIANVSLADTN  
S  
FT LAVVSTQLIMPQGEAGLQVPLIVGLILVLMVVLASLIYRRRLMKQDFSVPLPHSS  
S  
FT HMLRPRIFCSCPIGENSPILLSQQV\*  
FT 67..459  
FT /gene="pmel 17"  
FT /number=2  
FT 853..1014  
FT /gene="pmel 17"  
FT /number=3  
FT 1266..2396  
FT /gene="pmel 17"  
FT /number=4  
FT 2499..>2722  
FT /gene="pmel 17"  
FT /number=5  
FT SQ Sequence 2817 BP; 627 A; 772 C; 739 G; 679 T; 0 other;

Query Match 52.1%; Score 1132; DB 10; Length 2817;  
Best Local Similarity 99.9%; Pred. No. 0.00e+00;  
Matches 1133; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Db 1265 gaccaggtccttctcgtgagcgtgccagttcgccagttggatggagggaacaag 1324  
|||||  
Qy 669 GACCAGTGCCTTTCTCGTGAGCGTCTCCAGTTGGGGCTTGGATGAGGGAACAAG 728  
|||||  
Db 1325 cacttctgagaatcagcctctgaccttgccctccagctccatgccccagtgctat 1384  
|||||  
Qy 729 CACTTCTCAGAAATCAGCTCTGACCTTGGCTCCAGCTCCATGACCCAGCTGCTAT 788  
|||||  
Db 1385 ctgctgaagctgacctctctcacactgggactttggagacagtagtgaacccctgac 1444  
|||||  
Qy 789 CTGGCTGAAGCTGACCTCTCTACACCTGGGACTTTGGAGACAGTAGTGAACCTGATC 848  
|||||  
Db 1445 tctcgggacctgtggtcactcatactactggagcctggccagtcactgccagggtg 1504  
|||||  
Qy 849 TCTCGGCACTTGTGTCTACTATATCTTACCTGGAGCTGGCCAGTCACTGCCAGGTG 908  
|||||  
Db 1505 gtctcgaggtgcatctctcaactctgtggtctctccagttccagggaccaca 1564  
|||||  
Qy 909 GTCTCGAGGCTGGCACTCTCTCACTCTCTGGCTCTCCAGTTCCAGGCCACCACA 968  
|||||  
Db 1565 gatggcacaggccaactgcagagggccctaacaccagctggccaagtgcctactaca 1624  
|||||  
Qy 969 GATGGGCACAGCCCACTGCAGAGCCCTTACACACAGCTGGCCAACTGCTTACTACA 1028  
|||||  
Db 1625 gaagtgtgggtactacacctgggtcaggcgccaaactgcagagccctctggaaccacatct 1684  
|||||  
Qy 1029 GAAGTGTGGGTACTACACTGGTACAGGGCCAACTGCAGAGCCCTCTGGAACCACTCT 1088  
|||||

Db 1685 gtcaggtgccaaccactgaagtataaagcactgcacctgtgcagatgccaactgcagag 1744  
|||||  
Qy 1089 GTGACGTGGCCACCACTCAAGTCATAAGCACTGCACCTGTGCAGATGCCAACTGCAGAG 1148  
Db 1745 agcaaggtatgacacgtgaagaagtgccagttccagaggtcatggtaccacactgcca 1804  
|||||  
Qy 1149 AGCACAGGTATGACACCTGAGAAGGTGCCAGTTTCAGAGGTCTGGTATACCACCTGGCA 1208  
Db 1805 gagatgtcaactccagaggtcacaggtatgaacactgcagaggtatcaattggtgctt 1864  
|||||  
Qy 1209 GAGATGTCAACTCCAGAGGCTACAGGTATGACACCTGCAGAGGTATCAATTGTGTGCTT 1268  
Db 1865 tctggaaaccacagctgcacaggtacaactacagagtgggtggtagaccacagctagagag 1924  
|||||  
Qy 1269 TCTGGAAACCACAGCTGCACAGGTAACTACAGAGTGGGTGGACCAACACAGCTAGAGAG 1328  
Db 1925 ctacctatccctgagcctgaaggtccaagatgccagctcaatcatgtctacggaaagtatt 1984  
|||||  
Qy 1329 CTACCTATCCCTGAGCCTGAAGGTCAGAGTGCACAGCTCAATCATGTCTACCGAAGATTT 1388  
Db 1985 acaggttccctgggcccctgctggatggtacagcaccctaaagctggtggaagacaaa 2044  
|||||  
Qy 1389 ACAGGTTCCCTGGCCCCCTGCTGGATGGGTACAGCCACCTTAAAGCTGCTGAAGACAA 1448  
Db 2045 gtcccctggagttgtgtctgtatcgatagttccttttcogtcaacctggacattgtc 2104  
|||||  
Qy 1449 GTCCCCCTGGATTGTTCTGTATGATATGTTCTTTTCGTCACCCCTGGACATTGTC 1508  
Db 2105 cagggtattgaaagtcgcagagactctgcaggtgtgcgtccggtgaggggagtcattt 2164  
|||||  
Qy 1509 CAGGTTATTGAAGTGCACAGATCCTGCGAGGCTGTGCCGTGCGGTGAGGGGATCGATTT 1568  
Db 2165 gagctgactgtctctgccaaagcgggctgccaaagaaagcctgcattgagatctcatcg 2224  
|||||  
Qy 1569 GAGCTGACTGTGTCCTGCCAAGGGGGGCTGCCAAGGAAGCCCTGCATCGAGATCTCATCG 1628  
Db 2225 ccaggtgtccagccccctgccagcggctgtgccagcctgtgctaccagccagcctgc 2284  
|||||  
Qy 1629 CCAGGGTCCACAGCCCTGCCACGGCTGTGCCAGCTGTGTCAGCCAGCCAGCCTGC 1688  
Db 2285 cagctggttgcaccagactgaagggtgggtcggggacataactgctcctaagtgtct 2344  
|||||  
Qy 1689 CAGCTGTTCTGCACCAAGATGACTGAAGGTTGGCTCGGGGACATACTGCCTCAATGTGCT 1748  
Db 2345 ctggctgataccaacagctggcagtggtcagcaccagcttcatcgtcgtt 2398  
|||||  
Qy 1749 CTGGCTGATACCAACAGCCTGGCAGTGGTGCAGACCCAGCCTTATCATGCTGCTGGT 1802  
|||||

RESULT 7  
LOCUS BOVBRPE1A 1584 bp mRNA MAM 08-MAR-1993  
DEFINITION Bovine retinal pigment (RPE1) mRNA, 3' end.  
ACCESSION M81193  
NID gl62770  
KEYWORDS retinal pigment.  
SOURCE Bos taurus (Library: Unizap XR) calf retinal pigment epithelium  
cDNA to mRNA.  
ORGANISM Bos taurus  
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;  
Eutheria; Artiodactyla; Ruminantia; Pecora; Bovidae; Bovinae;  
Bovini.  
REFERENCE 1 (bases 1 to 1584)  
AUTHORS Kim, R.Y. and Wistow, G.J.

TITLE The cDNA RPE1 and monoclonal antibody HMB-50 define gene products preferentially expressed in retinal pigment epithelium  
JOURNAL Curr. Eye Res. 55, 657-662 (1992)  
COMMENT NCBI gi: 162770

FEATURES  
Location/Qualifiers

source  
1..1584  
/organism="Bos taurus"  
/dev\_stage="calf"  
/sequenced\_mol="cDNA to mRNA"  
/tissue\_type="retinal pigment epithelium"  
/tissue\_lib="Unizap XR"  
1..1477  
/partial  
/gene="RPE1"  
/note="ORF; NCBI gi: 162771"  
/codon\_start=2  
/product="retinal pigment"  
/db\_xref="PID:gl62771"

CDS

BASE COUNT 340 a 473 c 418 g 353 t  
ORIGIN  
Query Match 37.1%; Score 806; DB 33; Length 1584;  
Best Local Similarity 82.7%; Pred. No. 0.00e+00;  
Matches 1348; Conservative 0; Mismatches 230; Indels 52; Gaps 21;  
Db 1 ccaatactggcaagtctctggggggccagtgctcgaactgaacatcggaacagacaagcc 60  
Qy 509 CCAATACTGGCAATTTCTAGGGGGCCAGTGCTCTGGCTGAGCATTGGGACAGCGAGGC 568  
Db 61 aatgctgggcaacataaacatggaagtgaactgtctaccaccgcgggggtcccaagacta 120  
Qy 569 AATGCTGGGCACACACACCATGGAAATGACTCTCTACCATCGCGGGGATCCCGAGGCTA 628  
Db 121 tgtccctcctcactcactccagttcagcttcaccattactgaccagtgccctctctgt 180  
Qy 629 TGTCCCTCTTCTCATTCAGCTCAGCCTTCACCATTTACTGACGAGGTGCCTTTCGCT 688  
Db 181 gagtgtgtcctcagctcagggccttgatggaaggaacaagcctcctcgagaaagcagcc 240  
Qy 689 GAGGCTGCCAGTTGGGGGCTTGGATGAGGAGCAACAGCACTTCCTTGAGAAATCAGCC 748  
Db 241 tctgaccttgccctccagctccatgatcccaagtggctatttggtgggggtgaccttc 300  
Qy 749 TCTGACCTTTGCTCCCTCCAGCTCCATGACCCCACTGGCTATCTGGCTGAGCTGACCTCTC 808  
Db 301 ctacaactgggaatttggtgacagtacagggaacctgatctctcgggcaactcaagctcac 360  
Qy 809 CTACACCTGGGACTTTGGAGACAGTACTGGAACCTTGATCTCTCGGGCACTTGTGCTCAC 868  
Db 361 tcacacttacctagagtctggcccagtcactgcacaggtgggtgctgcagggtgccattcc 420  
Qy 869 TCATACTTACCTGGAGCCTGGGCCCACTGCTACTGCCACAGGTGGCTCTGCAGAGCTGCCATTC 928  
Db 421 tctcaactcctgtggctcctctccagttccagggcaactacagataggcatgtgacaactgc 480  
|||||



Qy	1889	GGCGAGCTTATGAGCAGACT-----TCTCCGTACCCCACTTGCACATAGACGACG	1942
Db	1396	ccagtggcctaagctcgcctggggtctccgctcttgcgccattggtagagcaaacccct	1455
Qy	1943	TCACGGGTGGGTACCCCGGCATCTTCTGCTTGTCCCATTTGGTGAACAGCCCCCT	2002
Db	1456	cctcagtgagacagagctcgaagtgccttatgtgaagtcatgatttaccagc-gtggac	1514
Qy	2003	CTTCAGTGGGACGAGGCTCAGTACTCTCATATGATGCTGTGATTTCTCGAGTTCAC	2122
Db	1515	agcaaggcctgtctttctctgtgttcccttcagagacactaccctgcctgaataaagac	1574
Qy	2063	AGAAACACCTATATTTCCCCAGCTCTCCCTGGGAGACTACTATTAACTGAATAAATAC	2122
Db	1575	tcagaacttg 1584	
Qy	2123	TCAGAGCGTG 2132	
RESULT	8		
LOCUS	MMU14133	1881 bp	mRNA
DEFINITION	Mus musculus pmel17 protein mRNA, complete cds.	ROD	07-JUL-1995
ACCESSION	U14133		
NID	9887940		
KEYWORDS	.		
SOURCE	mouse.		
ORGANISM	Mus musculus		
REFERENCE	1 (sites)		
AUTHORS	Kwon, B.S., Halaban, R., Ponnazhagan, S., Kim, K., Chintamaneni, C., Bennett, D. and Pickard, R.T.		
TITLE	Mouse silver mutation is caused by a single base insertion in the putative cytoplasmic domain of Pmel 17		
JOURNAL	Nucleic Acids Res. 23, 154-158 (1995)		
MEDLINE	95175358		
REFERENCE	2 (bases 1 to 1881)		
AUTHORS	Kwon, B.S.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-AUG-1994) Byoung S. Kwon, Indiana University School of Medicine, Microbiology and Immunology, 635 Barnhill Dr, Indianapolis, IN 46202, USA		
COMMENT	NCBI gi: 887940		
FEATURES	Location/Qualifiers		
source	1..1881		
	/clone="pmel17-10"		
	/strain="c57Bl/6"		
	/organism="Mus musculus"		
	/map="chromosome 10 si locus"		
	/chromosome="10"		
	/sex="female"		
	/cell_line="primary culture"		
	/cell_type="melanocyte"		
	/tissue_type="skin"		
	1..1881		
	/gene="pmel17"		
	/note="NCBI gi: 887941"		
	/codon_start=1		
	/product="pmel17 protein"		
	/db_xref="PID:g887941"		
	/translation="MVGQVRRSFLPVLVLSALLAVGALGSGRNODMIGVPRQLVTKTW		
	NRLQLPWFTEVQSGNCRGQGVSLRVINDGPTLVGANAFSLAHFPGSKVLPGDQV		
	IWANNNTINGSQWVGQPVYPQEPDDACVFPDGGCPSGGKPPFSEVYVWTKTWKTC		
CDS			





FT exon 95..229  
FT exon /number=4  
FT exon 623..781  
FT exon /number=5  
FT exon 1030..1755  
FT exon /number=6  
SQ Sequence 2001 BP; 457 A; 554 C; 492 G; 498 T; 0 other;  
  
Query Match 33.4%; Score 726; DB 10; Length 2001;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 1030 gaccaggtgccttctccgtgagcgtgtcccagttggggccttggatggagggaacaag 1089  
|||||  
Qy 669 GACCAGGTGCTTCTCCGTGAGGCTGCCAGTTCGGGCTTGGATGAGGGAACAAG 728  
|||||  
Db 1090 cacttctgagaatcagcctctgaccttggcctccagctccatgacccagtggtat 1149  
|||||  
Qy 729 CACTTCTGAGAAATCAGCCTCTCACCTTGGCTCCAGTCCATGACCCAGTGGCTAT 788  
|||||  
Db 1150 ctggtggaagctgacctctcctacacctgggactttggagacagtagtgaaccctgac 1209  
|||||  
Qy 789 CTGGCTGAAGCTGACCTCTCTCACCTGGGACTTTGGAGACAGTAGTGGAACTCTGATC 848  
|||||  
Db 1210 tctcgggcacttgtgtcaactcatactgagcctggcctggcctgacactgccaggtg 1269  
|||||  
Qy 849 TCTCGGCACTTGTGGTCACTACTACTTACCTTGGAGCTGGCCAGTCACTGCCAGGTG 908  
|||||  
Db 1270 gtectgaggctgccattctctcaactctgtggtcctcccagttgccagtcaccaca 1329  
|||||  
Qy 909 GTCTCGAGGCTGCCATTCCTCTCACTCTCTGGCTCTCCAGTTCAGGCGACACACA 968  
|||||  
Db 1330 gatggcacaggccaaactgcagagggccctaacaccacagctggccaaagtgcctactaca 1389  
|||||  
Qy 969 GATGGGCAAGGCAACTGACAGAGGCCCCACACACACAGCTGGCCAGTGGCTACTACA 1028  
|||||  
Db 1390 gaagttgtgggtactacacctgggtcaggcgccaaactgcagagccctctggaaccacatct 1449  
|||||  
Qy 1029 GAAGTTGTGGGTACTACCTCTGCTCAGGGCCAACTGCAGAGCCCTCTGGAACCACTCT 1088  
|||||  
Db 1450 gtccaggtgccaaaccactgaagtcataaagcactgcaactgtgcagatgccaaactgcagag 1509  
|||||  
Qy 1089 GTCCAGGTGCCAACCACTGAAGTCATAAGCACTGCACCTGCAGATGCCAACTGCAGAG 1148  
|||||  
Db 1510 agcacaggtatgacacctgagaaggtgccagtttcagaggtcatgggtaccacactggca 1569  
|||||  
Qy 1149 AGCACAGGTATGACACTGACGAGGTGCCAGTTTCAGAGGTATGGGTGACCACTGGCA 1208  
|||||  
Db 1570 gagatgtcaactccagaggtacaggtatgacacctgcagaggtatcaattgtggtcctt 1629  
|||||  
Qy 1209 GAGATGTCAACTCCAGAGCTACAGGTATGACACCTGCAGAGGTATCAATTGTGTGCTT 1268  
|||||  
Db 1630 tctggaaccacagctgcacaggttaacaactacagagtgggtggagaccacagctagagag 1689  
|||||  
Qy 1269 TCTGGAAACCAAGCTGCACAGGTAAACACTACAGAGTGGGTGGAGACCACTAGAGAG 1328  
|||||  
Db 1690 ctacctatccctgagcctgaaggtccagatgccaggtccaatcatgtctacggaaagtatt 1749  
|||||  
Qy 1329 CTACCTATCCCTGAGGCTGAAGTCCAGATGCCAGTCAATCATCTCTCGGAAGTATT 1388  
|||||  
Db 1750 acaggt 1755  
|||||  
Qy 1389 ACAGGT 1394  
|||||

RESULT 10  
LOCUS HSPMEL03 2001 bp DNA PRI 19-DEC-1995  
DEFINITION Human melanocyte protein Pmel 17 gene, exons 4, 5, and 6.  
ACCESSION U31797  
NID g1125059  
KEYWORDS .  
SEGMENT 3 of 5  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2001)  
AUTHORS Bailin, T., Lee, S.-T. and Spritz, R.A.  
TITLE Genomic organization and sequence of D12S53E (Pmel 17), the human homologue of the mouse silver (si) locus  
JOURNAL Unpublished (1995)  
REFERENCE 2 (bases 1 to 2001)  
AUTHORS Bailin, T.  
TITLE Direct Submission  
JOURNAL Submitted (17-JUL-1995) Richard A. Spritz, Medical Genetics, 445 Henry Mall, University of Wisconsin, Madison, WI 53706, USA  
COMMENT NCBI gi: 1125059  
FEATURES  
source  
1..2001  
/organism="Homo sapiens"  
/chromosome="12"  
/map="12q13-q14"  
/note="DSEG number: D12S53E"  
95..229  
/number=4  
623..781  
/number=5  
1030..1755  
/number=6  
BASE COUNT 457 a 554 c 492 g 498 t  
ORIGIN  
  
Query Match 33.4%; Score 726; DB 86; Length 2001;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 1030 gaccaggtgccttctccgtgagcgtgtcccagttggggccttggatggagggaacaag 1089  
|||||  
Qy 669 GACCAGGTGCTTCTCCGTGAGGCTGCCAGTTCGGGCTTGGATGAGGGAACAAG 728  
|||||  
Db 1090 cacttctgagaatcagcctctgaccttggcctccagctccatgacccagtggtat 1149  
|||||  
Qy 729 CACTTCTGAGAAATCAGCCTCTCACCTTGGCTCCAGTCCATGACCCAGTGGCTAT 788  
|||||  
Db 1150 ctggtggaagctgacctctcctacacctgggactttggagacagtagtgaaccctgac 1209  
|||||  
Qy 789 CTGGCTGAAGCTGACCTCTCTCACCTGGGACTTTGGAGACAGTAGTGGAACTCTGATC 848  
|||||  
Db 1210 tctcgggcacttgtgtcaactcatactgagcctggcctggcctgacactgccaggtg 1269  
|||||  
Qy 849 TCTCGGCACTTGTGGTCACTACTACTTACCTTGGAGCTGGCCAGTCACTGCCAGGTG 908  
|||||  
Db 1270 gtccagaggtgccattctctcaactctgtggtcctcccagttgccagtcaccaca 1329  
|||||  
Qy 909 GTCTCGAGGCTGCCATTCCTCTCACTCTCTGGCTCTCCAGTTCAGGCGACACACA 968  
|||||  
Db 1330 gatggcacaggccaaactgcagagggccctaacaccacagctggccaaagtgcctactaca 1389  
|||||

QY	969	GATGGGCACAGGCCAAC	TGCTACAGAGGCCCTTACAC	CACAGCTGGCCAGTGCCTACTACA	1028
Db	1390	gaagttgtggtaactacac	ctggtcagcgcccaactgcag	agccctctggaaaccacatct	1449
QY	1029	GAAGTTGTGGGTACAC	CTGTGTCAGCGCGCAACT	TCGAGAGCCCTCTGGAACACATCT	1088
Db	1450	gtgcaggtgccaaacac	gaagtcatagaacactgca	actgtgcagatgccaaactgcagag	1509
QY	1089	GTCCAGGTGCCAACAC	CTGAAGTCAATAGCACTG	CCACTGTGCAGATGCCAACTGCAGAG	1148
Db	1510	agcacagtgatgacac	ctggagaagtgccaagttc	csagaggtcatgggtaccacaactgcga	1569
QY	1149	AGCACAGTATGACAC	CTGAAAGGTGCCATTT	TCAGAGGTCATGGGTACCACTGGCA	1208
Db	1570	gagatgtcaactccag	aggtacagdtgacacctg	cagaggtatcaattgtgtgctt	1629
QY	1209	GAGATCTCACTCCAG	AGGCTACAGGTATGAC	ACTTCGACAGGTA	1268
Db	1630	tctggaaccacagctg	cacaggtgaacaactaca	gagtggtggagaccacagctagagag	1689
QY	1269	TCTGGACACACAGCT	GCACAGGTACACTAC	ACAGTGGGTGGAGACACACAGCTAGAGAG	1328
Db	1690	ctacctatccctgag	ctggaaggtccagatgcc	agctcaatctgtctacggaaagtatt	1749
QY	1329	CTACCTATCCCTGAG	CGCTGAAGTCCGATG	CCAGTCAATCATGTCTACGGAAAGTATT	1388
Db	1750	acaggt	1755		
QY	1389	ACAGGT	1394		

RESULT	11	
ID	HSMEI05	standard; DNA; PRI; 1178 BP.
AC	U31799;	
DT	23-DEC-1995 (Rel. 46, Created)	
DT	23-DEC-1995 (Rel. 46, Last updated, Version 1)	
DE	Human melanocyte protein Pmel 17 gene, exon 11 and partial cds.	
KW	*	
OS	Homo sapiens (human)	
OC	Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;	
OC	Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.	
RN	[1]	
RP	1-1178	
RA	Bailin T., Lee S.T., Spritz R.A.;	
RT	*Genomic organization and sequence of D12S53E (Pmel 17), the human	
RT	homologue of the mouse silver (si) locus*;	
RL	Unpublished.	
RN	{2}	
RP	1-1178	
RA	Bailin T.;	
RT	*	
RL	Submitted (17-JUL-1995) to the EMBL/GenBank/DBSJ databases.	
RL	Richard A. Spritz, Medical Genetics, 445 Henry Mall, University of	
RL	Wisconsin, Madison, WI 53706, USA	
CC	NCBI gi: 1125061	
Key	Location/Qualifiers	
FT	1..1178	
FT	/organism="Homo sapiens"	
FT	/chromosome="12"	
FT	/map="12ql3-q14"	
FT	/note="DSEG number: D12S53E"	
FT	join(U31808:158..232,U31807:128..241,U31807:400..540)	
FT	CDS	

FT FT :95..229,U31797:623..781,U31797:1030..1755,U31798:382..494  
FT FT U31798:736..821,U31798:932..1135,U31798:1238..1327,298..43  
2 )  
FT FT /notes="gp100; ME20; NCBI gi: 1125063"  
FT FT /codon\_start=1  
FT FT /product="melanocyte protein Pmel 17"  
FT FT /db\_xref="PID:gi1125063"  
FT FT /translation="MDLVIKRLLHLAVIGALLAVGATKVPKQNDMLGVSRLQLETKAW  
FT FT  
FT N RQLYPEWTEAQRLCDKRGQVSLKVSNDGPTLICANASFIALNFFCSQKVLPGQV1  
FT W VNNTINGSQVWGQVYPQETDADICFPDGGPCPGSGWSQKRSFYVWKTWGQYHQV  
FT L GCPVSGLSICTGRAMLGHTEVTVVHRRGSRSYVLAHSSSAFTITDQVPFVSYSQ  
FT FT RALDGGNKHFLRNQPLTFALQLHDPSCYLAEADLSYTNDFGDSGTLISRALVWTHITY  
FT L EPGVTAQVLIQAAILPSCGSSVPVCTDGHRETRAEAPNTACQVPTTEVWGTTCGQ  
FT FT PTAEPGCTTSVQVPTTEVISTAPVQMPAEASTGCTKPEVPVSEVMGCTTIAEMSTPEAT  
FT FT MTPAEVSIWLSGTTAAQVTTTENVETTARELPIPEPEGPDASSIMSTESITGSLGFL  
FT L DGTATLRLKQVPLDCLVLYRGSFVSVLIDVQGIESAEILQAVPSEGDFAFELTVSC  
FT Q GGLPKEACMEISSPGQPPAQLCQVLPFSEPAQCLVLHQILKGGSGTYCLNVSLADTN  
FT S LAVVSTQLIMPQGEAGLGQVPLVIGILLVMAVVLASLIYRRLRMKQDFSVPLPQHS  
FT S HWLRLPRTFCSCPIGENSPLLSGQV"  
SQ Sequence 1178 BP; 298 A; 231 C; 373 G; 276 T; 0 other;

Query Match	11.3%;	Score 245;	DB 10;	Length 1178;	
Best Local Similarity	99.6%;	Pred. No. 3.61e-110;			
Matches	246;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0
Db	295	aggcgagacttatgaagcaagacttctccgtaccocagttgccacatagacagcagtcac	354		
Qy	1887	AGGCGCAGACTTATGAAGCAAGACTTCCTCGTACCCAGTTGCCATAGCAGCAGTCAC	1946		
Db	355	tggtgtcgctaccocgcacatctctgtcttgtcccatgtgtgagaatagccccctctc	414		
Qy	1947	TGGCTGGGCTACCGCGCATCTCTGTCTCTGTGCCATTGGTCAGAACAGCCGCCCTCCTC	2006		
Db	415	agtgggcagcaggtctgagtactctcatatgatctgtgatttctcctggagttgcacaa	474		
Qy	2007	AGTGGGCAGCAGGCTCAGTACTCTCATATGATGCTGTGATTTCTCTGGAGTTGCAGAA	2066		
Db	475	acacatatcttcccocagctctccctgggagactactattaaactgaaataaataactcag	534		
Qy	2067	ACACCTATATTCCCCGAGCTCTTCCTCGGAGACTACTATTAACTGAAATAAATACTCAG	2126		
Db	535	agcctga	541		
Qy	2127	AGCCTGA	2133		
RESULT	12				
LOCUS	HSPMEL05	1178 bp	DNA	PRI	19-DEC-1995
DEFINITION	Human melanocyte protein Pmel 17 gene, exon 11 and partial cds.				
ACCESSION	U31799				
NID	g1125061				

Aug 27 08:56

US-08-231-565A-26.rgc

29

KEYWORDS  
SEGMENT 5 of 5  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1178)  
AUTHORS Bailin, T., Lee, S.-T. and Spritz, R.A.  
TITLE Genomic organization and sequence of D12553E (Pmel 17), the human  
homologue of the mouse silver (si) locus  
JOURNAL Unpublished (1995)  
AUTHORS 2 (bases 1 to 1178)  
Bailin, T.  
TITLE Direct Submission  
JOURNAL Submitted (17-JUL-1995) Richard A. Spritz, Medical Genetics, 445  
Henry Mall, University of Wisconsin, Madison, WI 53706, USA  
COMMENT NCBI gi: 1125061  
FEATURES  
source  
1..1178  
/organism="Homo sapiens"  
/chromosome="12"  
/map="12q13-q14"  
/note="DSEG number: D12553E"  
Join(U31808:158..232,U31807:128..241,U31807:400..546,  
U31797:95..229,U31797:623..781,U31797:1030..1755,  
U31798:382..496,U31798:736..821,U31798:932..1135,  
U31798:1238..1327,298..432)  
/note="gpl00; ME20; NCBI gi: 1125063"  
/codon\_start=1  
/product="melanocyte protein Pmel 17"  
/db\_xref="PID:g1125063"  
/translation="MDLVKRLHLHNAVIGALLAVGATKVRNQMLGVSRLRTKAW  
NRQYPTWTEAQRLDQMRGQVSLKNSDPTLIGANNAFSIALNFFGSRVLPDQV  
IWNQITNGSQVWGQVPYPOETDDACIFPDGPGCPGSGMSOKRSFYVVKWTGQVY  
QVLGGPVSGLSIGTCRMLGHTMEVTVHRRGSRVYPLHSSASATIDQVPSVS  
VSQRLADGKNHLELRNQPLTFALQLDHPGTYLAEDLSYTWDFGDSGTLISRALV  
THYLEPGEVTAQVVLQAAIPLTSCGSPVPCTDGHRTAEAPNTTAGQVPTTEWG  
TTPQAPTAEP SGTSVQVPTTEVI STAPVQMP TAESTCMTEKVPVSEVMTTIAEM  
STPEATGMPEAIVSVLSGTTAAQVTTTEWVETARELP IPEEGPDASSIMTESI  
TGS LGP LLDGTATLRVLRQVPLDCLVRYGFSVTIDIVOGIESAIIQAVP SGED  
AFELTVSCQGLPKAEIMEISPCOPPAQRLCQVLP SPACQLVLIHQILKGGSTYC  
LNVSLADTNSLVVSTQLIMP GQEAGLGQVPLIVGILLVIMAVLASLIYRRLMKQD  
FSVPQLPHSSHMLRIFRCSPIGENSPILLSQQV"

CDS

1..1559  
/organism="Homo sapiens"  
/chromosome="12"  
/map="12q13-q14"  
/note="DSEG number: D12553E"  
382..496  
/number=7  
736..821  
/number=8  
932..1135  
/number=9  
1238..1327  
/number=10  
BASE COUNT 298 a 231 c 373 g 276 t  
ORIGIN

Query Match 11.3%; Score 245; DB 86; Length 1178;  
Best Local Similarity 99.6%; Pred. No. 3.61e-110;  
Matches 246; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Db 295 aggcagactatgaagaagacttccgtaccacagttgccatagcagcagtcac 354  
|||||  
Qy 1887 AGGCGAGACTTATGAAGAAGACTTCCCTCCGACCCAGTTGCCATAGCAGTCAC 1946  
|||||  
Db 355 tggctgcgtaccacagcattcttctgtcttctgtccattggtgagaatagccctctc 414  
|||||  
Qy 1947 TGGCTGCGTACCCCGCATCTTCTGCTCTTCTGCTGCTGCTGCTGCTGCTC 2006  
|||||  
Db 415 agtgggacagcgtctgagtaactctcatatgatgctgatttctcggagttgacagaa 474  
|||||  
Qy 2007 AGTGGGACGAGGTCTGACTACTCTCATATGATGCTGATTTCTCGAGTTCACAGAA 2066  
|||||  
Db 475 acacttatattccccccagctcttccctggagactactaactgaataaatactcag 534  
|||||

Aug 27 08:56

US-08-231-565A-26.rgc

30

Qy 2067 ACACCTATATTCCCGCAGCTCTCCCTGGGAGACTACTATTAACTGAATAAATCTCAG 2126  
Db 535 agcctga 541  
|||||  
Qy 2127 AGCCTGA 2133  
|||||  
RESULT 13  
LOCUS HSPMEL04 1559 bp DNA PRI 19-DEC-1995  
DEFINITION Human melanocyte protein Pmel 17 gene, exons 7, 8, 9, and 10.  
ACCESSION U31798  
NID g1125060  
KEYWORDS  
SEGMENT 4 of 5  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1559)  
AUTHORS Bailin, T., Lee, S.-T. and Spritz, R.A.  
TITLE Genomic organization and sequence of D12553E (Pmel 17), the human  
homologue of the mouse silver (si) locus  
JOURNAL Unpublished (1995)  
REFERENCE 2 (bases 1 to 1559)  
AUTHORS Bailin, T.  
TITLE Direct Submission  
JOURNAL Submitted (17-JUL-1995) Richard A. Spritz, Medical Genetics, 445  
Henry Mall, University of Wisconsin, Madison, WI 53706, USA  
COMMENT NCBI gi: 1125060  
FEATURES  
source  
1..1559  
/organism="Homo sapiens"  
/chromosome="12"  
/map="12q13-q14"  
/note="DSEG number: D12553E"  
382..496  
/number=7  
736..821  
/number=8  
932..1135  
/number=9  
1238..1327  
/number=10  
BASE COUNT 340 a 392 c 413 g 414 t  
ORIGIN

Query Match 9.6%; Score 209; DB 86; Length 1559;  
Best Local Similarity 100.0%; Pred. No. 1.98e-90;  
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 930 ggcgtgccaaagaagcctgcatggagatctcatgccagggtgccagccctgccagc 989  
|||||  
Qy 1594 GCGTCCCAAGAGAGAGCTGCATGGAGATCTCATGCCAGGGTGCCAGCCCTGCCAGC 1653  
|||||  
Db 990 ggcgtgccagcctggtaccagccagcagcagcagcagcagcagcagcagcagcagc 1049  
|||||  
Qy 1654 GCGTGTGCCAGCCTGTGTACCCAGCCAGCCTGCCAGCTGCCAGCTTCTGCACGATCTGA 1713  
|||||  
Db 1050 aggtggtcggggacactgctcaatgtgtctgtgctgataccaacagcctggcag 1109  
|||||  
Qy 1714 AGGGTGGCTCGGGGACATACTGCCTCAATGTCTCTCTGGCTGATACCAACAGCCTGGCAG 1773  
|||||  
Db 1110 tggtcagcagccagcagcttatcagcctggt 1138  
|||||

QY 1774 TGGTCAGCAGCCAGCTTATCATGCTGGT 1802

RESULT 14  
ID HSMEL04 standard; DNA; PRI; 1559 BP.  
AC U31798;  
DT 23-DEC-1995 (Rel. 46, Created)  
DT 23-DEC-1995 (Rel. 46, Last updated, Version 1)  
DE Human melanocyte protein Pmel 17 gene, exons 7, 8, 9, and 10.  
KW  
OS Homo sapiens (human)  
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;  
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.  
RN [1]  
RP 1-1559  
RA Bailin T., Lee S.T., Spritz R.A.;  
RT "Genomic organization and sequence of D12S53E (Pmel 17), the human  
RT homologue of the mouse silver (si) locus";  
RL Unpublished.  
RN [2]  
RP 1-1559  
RA Bailin T.;  
RT ?  
RL Submitted (17-JUL-1995) to the EMBL/GenBank/DBJ databases.  
RL Richard A. Spritz, Medical Genetics, 445 Henry Mall, University of  
RL Wisconsin, Madison, WI 53706, USA  
CC NCBI gi: 1125060  
FH Key Location/Qualifiers  
FT source 1..1559  
FT /organism="Homo sapiens"  
FT /chromosome="12"  
FT /map="12q13-q14"  
FT /note="DSEG number: D12S53E"  
FT exon 382..496  
FT /number=7  
FT exon 736..821  
FT /number=8  
FT exon 932..1135  
FT /number=9  
FT exon 1238..1327  
FT /number=10  
SQ Sequence 1559 BP; 340 A; 392 C; 413 G; 414 T; 0 other;

Query Match 9.6%; Score 209; DB 10; Length 1559;  
Best Local Similarity 100.0%; Pred. No. 1.98e-90;  
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 930 ggctgccaggagcctgtgctaccagccagcctgccagctgtctgtcacagatactga 989  
|||||  
QY 1594 GCCTGCCAGGAGAGCTGTGATGAGATCTCATGCCAGGCTGCCAGCCCTGCCAGC 1653  
Db 990 ggctgtgccagctgtgctaccagccagcctgccagctgtctgtcacagatactga 1049  
|||||  
QY 1654 GGCTGTGCCAGCCTGTCTACCCAGCCAGCCTGCCAGCTGTCTGCCAGCAGTACTGA 1713  
Db 1050 aggggtggctcggggagacatactcctcaatgtgtctgtgtgataccagagcctggcag 1109  
|||||  
QY 1714 AGGGTGGCTCGGGGACATACCTGCTCAATGTCTGTGGCTGATACCAAGAGCTGGCAG 1773  
Db 1110 tggtcagcagccagccttatcatgcctggt 1138  
|||||  
QY 1774 TGGTCAGCAGCCAGCTTATCATGCTGGT 1802

RESULT 15  
LOCUS HSPMEL02 673 bp DNA PRI 19-DEC-1995  
DEFINITION Human melanocyte protein Pmel 17 gene, exons 2 and 3.  
ACCESSION U31807  
NID g1125058  
KEYWORDS  
SEGMENT 2 of 5  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; mitochondria eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 673)  
AUTHORS Bailin, T., Lee, S.-T. and Spritz, R.A.  
TITLE Genomic organization and sequence of D12S53E (Pmel 17), the human  
TITLE homologue of the mouse silver (si) locus  
JOURNAL Unpublished (1995)  
REFERENCE 2 (bases 1 to 673)  
AUTHORS Spritz, R.A.  
TITLE Direct Submission  
JOURNAL Submitted (17-JUL-1995) Richard A. Spritz, Medical Genetics, 445  
Henry Mall, University of Wisconsin, Madison, WI 53706, USA  
COMMENT NCBI gi: 1125058  
FEATURES  
source  
1..673  
/organism="Homo sapiens"  
/chromosome="12"  
/map="12q13-q14"  
/note="DSEG number: D12S53E"  
128..241  
/number=2  
400..546  
/number=3  
BASE COUNT 186 a 142 c 179 g 166 t  
ORIGIN

Query Match 6.9%; Score 150; DB 86; Length 673;  
Best Local Similarity 100.0%; Pred. No. 1.46e-58;  
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 396 aggtggcaggtgtccctcaggtcagtaagtggtggcctacactgattgtgcaaatgc 455  
|||||  
QY 224 AGGTGGTCAAGTGTCCCTCAAGGTCAGTAATGATGGGCTACACTGATTGGTCAATGC 283  
Db 456 ctctctctctgtgctgaacttcctcctggaagcgaagaaagattatgccagatggcaggt 515  
|||||  
QY 284 CTCCTTCTCTATTGGCTTGAATCTCCCTGGAAGCCAAAGGTATTGGCCAGATGGCAGGT 343  
Db 516 tatctgggtcaacaataccatcatcaatgg 545  
|||||  
QY 344 TATCTGGGTCAACATATACCATCATATGG 373

Search completed: Tue Aug 27 09:09:11 1996  
Job time : 1535 secs.

Aug 27 09:00

US-08-231-565A-26.mg

1

\*\*\*\*\*

MPERCH\_NN n.a. - n.a. database search, using Smith-Waterman algorithm

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

Run on: Tue Aug 27 09:09:30 1996; MasPar time 177.89 Seconds

Tabular output not generated.  
868.381 Million cell updates/sec

Title: >US-08-231-565A-26  
Description: (1-2172) from US08231565A.seq  
Perfect Score: 2172  
N.A. Sequence: 1 GTCCAGCGCCATTACCAATC.....AAAAAAAAAAAAAAAAAAAA 2172  
Comp: CAGCTCGCGGTATGCTTAG.....TTTTTTTTTTTTTTTTTTTT

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 91712 seqs, 35561689 bases x 2

Post-processing: Minimum Match 04  
Listing first 45 summaries

Database: n-gene23

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17

Statistics: Mean 9.903; Variance 7.604; scale 1.302

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2172	100.0	2172	17	T02716	MART-1 melanoma antig 0.00e+00
2	2101	96.7	2115	16	Q96055	Sequence encoding mel 0.00e+00
3	1970	90.7	2131	17	T03760	Melanoma-specific imm 0.00e+00
4	88	4.1	1047	2	Q10572	Human Natriuretic Pep 2.95e-29
5	86	4.0	1047	2	Q10572	Human Natriuretic Pep 3.04e-28
6	45	2.1	458	15	Q90525	Rat SIII 15 kDa subun 2.36e-08
7	44	2.0	204	1	N81164	Base substituted E.co 6.69e-08
8	43	2.0	350	16	Q99380	Rat allograft inflam 1.88e-07
9	43	2.0	1338	11	Q65607	Rabbit zona pellucida 1.88e-07

Aug 27 09:00

US-08-231-565A-26.mg

2

10	44	2.0	1710	8	Q47966	Rape acyl-ACP thioest 6.69e-08
11	43	2.0	1813	15	Q88760	Human ubiquitous nucl 1.88e-07
12	42	1.9	91	9	Q51746	Oligonucleotide probe 5.27e-07
13	42	1.9	91	9	Q51746	Oligonucleotide probe 5.27e-07
14	41	1.9	699	10	Q58663	Human superoxide dis 1.46e-06
15	41	1.9	756	12	Q73729	Rape abscission/dehis 1.46e-06
16	42	1.9	861	17	Q99552	Thrombopoietin coding 5.27e-07
17	41	1.9	971	13	Q73760	RP-8 Programmed cell 1.46e-06
18	41	1.9	1325	5	Q31873	Cyclin D1 gene. 1.46e-06
19	41	1.9	1325	5	Q31873	Human cyclin D1. 1.46e-06
20	41	1.9	1394	1	Q05879	Placenta-specific pro mPD-1 cDNA. 1.46e-06
21	41	1.9	1972	16	Q95309	Murine programmed cel 1.46e-06
22	41	1.9	1973	10	Q55770	Human pro-urokinase c 1.46e-06
23	41	1.9	2377	3	Q20360	Human c-myc far upstr 5.27e-07
24	42	1.9	2381	12	Q68910	Human c-myc far upstr 5.27e-07
25	42	1.9	2384	12	Q68909	Human pro-urokinase c 1.46e-06
26	41	1.9	2427	1	Q04107	Plasmid pAU157 sequen 1.46e-06
27	41	1.9	2696	2	N60889	Sequence of plasmid p 1.46e-06
28	41	1.9	2696	2	N60889	Shuttle vector pMW16 1.46e-06
29	41	1.9	2697	2	N60864	Clone Imd2 encoding i 5.27e-07
30	41	1.9	3138	2	Q11712	Murine glycosylation 4.04e-06
31	42	1.9	3504	1	Q05304	Hybrid human/pig SOD. 4.04e-06
32	40	1.8	635	13	Q75445	Human SOD. 4.04e-06
33	40	1.8	667	11	Q61575	P14(T) allergen clone 4.04e-06
34	40	1.8	688	11	Q61572	Fragment of plasmid p 4.04e-06
35	40	1.8	700	3	Q22352	CAD cDNA isolated fro 4.04e-06
36	40	1.8	887	1	Q05160	Human anti-HBs heavy 4.04e-06
37	40	1.8	1393	6	Q38027	Tomato acid invertase 4.04e-06
38	40	1.8	1576	8	Q49944	Human heregulin-beta2 4.04e-06
39	40	1.8	2431	5	Q31546	Full-length foetal hu 4.04e-06
40	40	1.8	2431	5	Q31546	P-meta-1 metastasis-s 4.04e-06
41	40	1.8	2445	7	Q48393	pmx34 cDNA insert. 4.04e-06
42	40	1.8	3207	3	Q14263	Plasmid sequence enco 4.04e-06
43	40	1.8	3312	2	N71034	pl50.95 alpha subunit 4.04e-06
44	40	1.8	3628	2	N60848	
45	40	1.8	4589	1	Q06068	

ALIGNMENTS

RESULT 1	
ID T02716	standard; cDNA; 2172 BP.
AC T02716;	
DT 20-APR-1996	(first entry)
DE MART-1 melanoma antigen cDNA25.	
KW cDNA25; MART-1; melanoma antigen recognised by T-cells;	
KW gp100 antigen derivative; melanoma; metastatic melanoma;	
KW tumour-associated antigen; immunogen; diagnosis; prognosis;	
KW prophylaxis; therapy; vaccine; ds.	
OS Mammalian.	
FH Key	Location/Qualifiers
FT CDS	38..2038
FT /*tag= a	
FT /note= "cDNA25 melanoma antigen"	
PN W09529193-A2.	
PD 02-NOV-1995.	
PF 21-APR-1995; U05063.	
PR 22-APR-1994; U0521565.	
PR 05-APR-1995; US-417174.	
PA (USSH ) US SEC DEPT HEALTH.	
PI Kawakami Y, Rosenberg SA;	
DR WPI; 95-382963/49.	
DR P-PSDB; R84854.	
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also	

PT vectors, host cells and antibodies, used to detect, treat and immunise animal against melanoma.

PS Disclosure; Fig 4A-4B; 184pp; English.

CC The nucleic acid encodes cDNA25, a melanoma antigen (MART-1) which is recognized by T-lymphocytes. cDNA25 is a derivative of the melanocyte-melanoma-specific antigen gp100 (see R84855).

CC Antigen cDNA25 is a source of immunogenic peptides (see R84199)

CC which are optionally modified (see R84200-R84211) and used in medicaments, especially vaccines, for the treatment or prevention (by immunization) of melanoma. Antibodies against cDNA25 and its immunogenic peptides may be used in the detection and isolation of the antigen from a sample, the detection of which is indicative of a disease state (melanoma or metastatic melanoma).

SQ Sequence 2172 BP; 512 A; 594 G; 578 G; 488 T;

Query Match 100.0%; Score 2172; DB 17; Length 2172;

Best Local Similarity 100.0%; Pred. No. 0.00e+00;

Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 gtcagcgccattaccatcgaccggggaagacacacatggtctggtcctaaagat 60

|||||

Qy 1 CTCGACGGCCATTACCAATCGGCNCGGGGAGAACACATGGATCTGGTGTAAAGAT 60

Db 61 gctcttctcatttgctgatagtgcttgcgtggtggtgggtctacaaaagtacca 120

|||||

Qy 61 GCCTTCTTCATTGCTGTGATAGTGTCTTGTGCTGTGGGGGCTCAAAAGTATCCCA 120

Db 121 gaaacaggactgcttggctctcaagcgaactcagaacaaagcctggaacagcagc 180

|||||

Qy 121 GAAACAGGACTGGTGGTGTCTCAGGCAACTCAGAACCAAGGCTTGAACAGGCAGC 180

Db 181 tctatccagagtgacacagaagccagagacttgactgctggagagtggtcgaagtccc 240

|||||

Qy 181 TGTATCCAGAGTGGACAGAGCCAGAGACTTGACTGCTGGAGAGTGTCAAGTGTCCC 240

Db 241 tcaaggtcagtaagtggcctacactgattggtgcaaatgcctctctctattgctt 300

|||||

Qy 241 TCAGGTCAGTAATGATGGGCCCTACACTGATTTGGTGCAATTCCTCTCTATTGGCT 300

Db 301 tgaacttcctggagaccaaaggtattgccagatggcagagttactcgtgggtcaacaata 360

|||||

Qy 301 TGAACCTTCCGTGAAGCCAAAGGATTTGCCAGATGGCGAGGTTATCTGGGTCAACAATA 360

Db 361 ccaatcacaatgggagccaggtgtgggagggacagccagtgatccccaggaactgaag 420

|||||

Qy 361 CCATCATCAATGGGAGCCAGGTGTGGGAGGACAGCCAGTGTATCCCAGGAAACTGAGG 420

Db 421 atgctgcattctccctgattggtgacactggcccatctggctcttggctcagaagaaa 480

|||||

Qy 421 ATGCCCTGCATCTCTCCGTATGTGTGACNCTTGGCCATCTGGCTCTTGGTCTCAGAGAGAA 480

Db 481 gctttgtttatgctcgtgaagacactgggccaactactggcaatttctagggggccagtg 540

|||||

Qy 481 GCTTTGTTTATGTCTGGAGAACCTGGGGCCCAATACTTGGCAATTTCTAGGGGGCCAGTGT 540

Db 541 ctgggctgaacattggagacggcaggccaactgctgggcacacacacacatggaagtga 600

|||||

Qy 541 CTGGGCTGACCAATTGGGACGGCAGGGCAATGCTGGGCACACACACCAATGGAAGTGA 600

Db 601 tctaccatcgccggggtatccggagctatgtgcctcttgcctcattccagctcagccttca 660

|||||

Qy 601 TCTACCATCGCCGGGATATCCCGAGACTATGTGCTCTTGTCTTCAATCCAGCTCAGCCTTCA 660

Db 661 ccattactgaccaggtgctcttctcgttgagcgtgtcccagttgcgggcttggatggag 720

Qy 661 CCATTACTGCACAGGTGCTTTCTCGGTGACGGGTGCCAGTTGCGGGGCTTGGATGGAG 720

Db 721 ggaacaagcacttctctgagaatcagcctctgcactttggcctcagctccatgacccca 780

|||||

Qy 721 GGAACAAGCACTTCTCTGAGAAATCAAGCTCTGACCTTTGGCCCTCAGCTCCATGACCCCA 780

Db 781 gtgctatctggctgaagctgacctctctacacctgggaactttggadacagtagtggaa 840

|||||

Qy 781 GTGGTATCTGGCTGAAGCTGACCTCTCTACACCTTGGGACTTTGGAGACAGTAGTGGAA 840

Db 841 cctgatctctcgggcacttggtcactcatacttacctggagcctggcccaagtcaetg 900

|||||

Qy 841 CCTGATCTCTGGGCACTTGTGTGCTACTTACTTACCTTGGAGGCTTGGCCCACTCACTG 900

Db 901 cccagtggtcctgcaggctgcccattcctctcaactcctgtggctcctccccagttccag 960

|||||

Qy 901 CCCAGGTGCTCTGCAGGCTGCCATTCCTCTCACCTCTCTGTGGCTCTCCCAAGTTCCAG 960

Db 961 gcaaccacagatgggcacaggccaaactgcagaggccctaacacacacagctggccaaagtgc 1020

|||||

Qy 961 GCACCACAGATGGGCACAGGCCAACTGGCAGAGGGCCCTAACACCAACAGCTGGCCAAAGTGC 1020

Db 1021 ctactcacagagttggtggtactacacctggtcaggcgccaaactgcagagccctctggaa 1080

|||||

Qy 1021 CTACTACAGAAAGTTGTGGGTACTACACTGCTGGCGGCCAACTGCAGAGGCCCTCTGGAA 1080

Db 1081 ccaatctgtgcaggtgccaaacctgaagtcatagcaactgcaacctgtgcagatgccaa 1140

|||||

Qy 1081 CCACATCTGTGCAGGTGCCAACCACTGAAGTCAATAAGCACTGCACTGTGCAGATGCCAA 1140

Db 1141 ctgcagagacacaggtatgacacctgagaagtgccaggtttcagaggtcatgggtacca 1200

|||||

Qy 1141 CTCGACAGACACAGGTATGACACTGAGAAGGTGCCAGTTTTCAGAGGTCTATGGTACCA 1200

Db 1201 cactggcagagatgtcaactccagaggctacaggtatgacacctgcagagtgatcaattg 1260

|||||

Qy 1201 CACTGGCAGAGATGTCAACTCCAGAGGCTACAGGTATGACACCTGCAGAGGTATCAATTG 1260

Db 1261 tggctcttctggaaccacagctgcacaggttaacaaactacagagtggtggagaccacag 1320

|||||

Qy 1261 TGGTCTCTTCTGCAACCAACAGCTGCCAGGTTAAACACTACAGAGTGGGTGGACACCAAG 1320

Db 1321 ctgagagagctaacatccctgcagcctgaaggtccagatgcccagctcaatcatgtctacgg 1380

|||||

Qy 1321 CTAGAGAGCTACCTATCCCTGAGCCTGAAGTCCAGTCCAGTCAATCATGTCTACGG 1380

Db 1381 aaagtattacaggttccctgggccccctgctggatgggtacagccaccttaaggctggtga 1440

|||||

Qy 1381 AAGTATTACAGTTCCCTGGGCCCCCTGCTGCTGATGCTAGACGCCACCTTAAGGCTGTGA 1440

Db 1441 agagacaagtccccctggattgttctgtatcgatatggttctctttccgtcacctgg 1500

|||||

Qy 1441 AGAGCAAGTCCCCCTGGATGTGTTCTGTATCGATATGTGTTCTTTTCCCTCACCTGG 1500

Db 1501 acattgtccaggggtattgaaagtgccagatcctgcagcctgtgccgtccgggtgaggggg 1560

|||||

Qy 1501 ACATTGTCCAGGGTATTGAAAGTGGCCGAGATCTCTGCAGGCTGTGCCGCTCAGGGGG 1560

Db 1561 atgcatttgagctgactgtcctgccaagggcggtgctcccaaggaagcctgcattggaga 1620

|||||

Qy 1561 ATGCATTTGAGCTGACTGTCTCTGCCAAGGGGGCTGCCAAGGAAGCAAGCTGTCATGGAGA 1620

Db 1621 tctcatcgccaggggtgcagggccccctgcccagcgctgtgccagcctgtgctaccagcc 1680

```
Qy 1621 TCTCATGCCAGGTCGCCAGCCCTGCCAGCGGCTGTCAGCCTGTGCTACCCAGCC 1680
Db 1681 cagcctgccagctggttctgaccagatactgaaggggtgctcggggacatactgctca 1740
Qy 1681 CAGCCTGCCAGCTGCTTCTGCACAGATACTGAAGGGTGGCTCGGGGACATACTGGCTCA 1740
Db 1741 atgtctctggctgatacaacagcctggcagtggtgagcaccagcttcatcatcctg 1800
Qy 1741 ATGTGCTCTGGCTGATACCAACAGCCTGGCAGTGTGAGCAGCCAGCTTATCATGCCTG 1800
Db 1801 qcaagaagcagccttggcaggttccgctgctgctggcagctcttggctggtggtgatgg 1860
Qy 1801 CTCAGAAGCAGCCCTTGGCCAGGTTCCGCTGATCGTGGCATCTTGGTGGTGTTCATGG 1860
Db 1861 ctgtggtcctgcatctctgatataagcgcagacttatgaagaagactttccgtac 1920
Qy 1861 CTGTGGTCTTGCATCTCTGATATATAGCCGACACTTATCAAGCAAGACTTCTCGGTAC 1920
Db 1921 cccagttgccacatgacagcagctcactggtgcgtctaccccgcactcttctgtctgtc 1980
Qy 1921 CCAGTTGCCACATAGCAGAGTCACTGGCTGGCTACCCGCCATCTTCTGCTCTGCTC 1980
Db 1981 ccattggtgagaaagcccccctcctcagtggcagcagctgctgactctcatatgatg 2040
Qy 1981 CCATTGGTGAACAGCCCTCCTCAGTGGGCGAGGCTCTGAGTACTCTCATATGATG 2040
Db 2041 ctgtgatttctcgtgagtgagaaacactatattccccagcttctccctggagagac 2100
Qy 2041 CTGTGATTTCTCGAGTGCACAGAAACACCTATATTTCCCCAGTCTTCCCTGGCAGAC 2100
Db 2101 tactattaactgaaataaataactcagagcctgaaaaaaataaaaaaaataaaaaaa 2160
Qy 2101 TACTATTAACGTGAATAACTCAGAGCCTGAATAAATAAATAAATAAATAAATAA 2160
Db 2161 aaaaaaaataa 2172
Qy 2161 AAAAAAATAA 2172

RESULT 2
ID Q96055 standard; cDNA to mRNA; 2115 BP.
AC Q96055;
DE Sequence encoding melanoma associated antigen gp100.
KW Melanoma; antigen; vaccine; immunogen; primer; probe; detection;
KW identification; tumour; gp100; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS
FT /*tag= a
FT /product= Melanoma associated antigen gp100.
FT misc_signal 1..81
FT /*tag= b
FT misc_feature 1792..1870
FT /*tag= c
FT /label= Transmembrane domain.
FT misc_binding 262..264
FT /*tag= d
FT /bound_moeity= Carbohydrate.
FT misc_binding 337..339
FT /*tag= e
FT misc_binding 352..354
FT /*tag= f
```

```
FT misc_binding 982..984
FT /*tag= g
FT misc_binding 1723..1725
FT /*tag= h
PN EP-668350-A1.
PF 23-AUG-1995.
PD 14-FEB-1995; 200348.
PR 16-FEB-1994; EP-200337.
PR 21-DEC-1994; EP-203709.
PA (ALKU ) AKZO NOBEL NV.
PI Adema GU, Flgdon CG;
DR WPI; 95-284790/38.
P-PSDB; R78646.
PT Melanoma associated antigen gp100 - used in vaccines and for the
PS detection of tumours
PS Claim 2; Page 19-22; 40pp; English.
CC Immunogenic peptides derived from the melanoma associated antigen
CC may be used in the production of vaccines. Nucleotide sequences
CC encoding the immunogenic peptides may be used as primers and probes
CC in the detection of melanoma cells. Tumour infiltrating lymphocytes
CC capable of binding to the melanoma associated antigen can be
CC cultured ex vivo and returned to melanoma particles, and when
CC radiolabelled, they may be used to identify tumour deposits.
SQ Sequence 2115 BP; 469 A; 587 C; 575 G; 484 T;

Query Match 96.7%; Score 2101; DB 16; Length 2115;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 2103; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 11 ggaagaacacaaatggatgctggtctaaaaagatgcttcttcttcttggctgtagagtg 70
Qy 28 GGAAGAACACAAATGGATCTGGTCTGCTTAAAGATGCTTCTTCATTTGGCTGTGATAGTG 87
Db 71 ctttctgctgtggggctacaaaagtaccagaaacacagagctgcttggctgtctcaaa 130
Qy 88 CTTTCTGCTGTGGGGCTTACAAAGTACCAGAAACACGAGACTGGCTTGGTGTCTCAA 147
Db 131 ggcactcagaacaaacgctcctcctcctcctcctcctcctcctcctcctcctcctcctc 190
Qy 148 GGCMACTCAGAACCAAGCCTGGAAACAGGAGCTGTATCCAGAGTGGAGAGCCACAG 207
Db 191 gacttgactgctggagaggtggctcaagtgtccctcaagtgcagtgcagtgcagtgcacac 250
Qy 208 GACTTGACTGCTGGAGAGGTGGTCAAGTGTCCCTCAAGGTCAGTATGATGGGCTTACAC 267
Db 251 tgattggtgcaaatgctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 310
Qy 268 TGATTGGTGCATAATGCCTCCTCTCTATTGGCTTGAAGTTCCTTGGAGAGCCCAAGGTAT 327
Db 311 tgccagatgggaggttatctctgggtcaacaataaccatcatcaatggagccaggtgtggg 370
Qy 328 TGCAGATGGGCAGGTTATCTGGGTCAACAATATCATCATATGGGAGCCAGGCTGTGGG 387
Db 371 gaggacagccagtgatccccaggaagaaactgacatgcctgcctcctcctcctcctcctcct 430
Qy 388 GAGGACGCCAGCTGATCCCCAGGAAACTGACGATGCCCTGCATCTTCCCTCATGGTGGAC 447
Db 431 cttgccatctggctcttggctctcagaagaagagcttggttttatgtctggaagacctggg 490
Qy 448 CTTGCCCATCTGGCTCTTGGTCTCAGAGAGAGAGCTTTGTTTATGCTTGGAGACCTGGG 507
Db 491 qccaataactggcaagtcttagggggccagctgtctgggctgagcattgggacagcgaggg 550
Qy 508 GCCAATACTGGCAATTTCTAGGGGGCCCGAGCTGTCTGGGCTGAGCATTTGGGACAGCGAGG 567
```



Db	551	caatgctgggcacacacacacacattgaagagtgaactgtaccatcgccgggagctcccgagct	610
Qy	568	CAATGCTGGGCACACACACCAATGGANGTCACTGTCTACCATCGCGGGGATCCGGAGCT	627
	611	atgtgctcttgcctcattccagctcagccttcaccattactgaccaggtgcctttctccq	670
Qy	628	ATGCGCTCTTGCTCAATTCAGCTCAGCTCTCCACATFACACAGGTGCTTTCGCG	687
	671	tgagcgtgtccagttgcggggccttggatggagggaaacaagcacttccctgagaatacag	730
Qy	688	TGAGGCTGTCCCAAGTTGGGGGCCCTTGGATGGAGGGAAACAAGCACTTCTCGTGAANAATCAG	747
	731	ctctgaaccttgccctccagctccatgaacccagtgagctatctgctgaagctgaacct	790
Qy	748	CTTGACCTTTGCCCTCCAGCTCCATGACCACGGTGGCTATCTGGCTGAAGCTGACCTCT	807
	791	ctctaacacctggagctttggagacagtagtgaacacctgatctctcgggcaaccttgggtca	850
Qy	808	CCTACACCTGGGACTTTGGAGACAGTAGTGGNAACCTGATCTCTCGGGCACTTGTGTCA	867
	851	ctcatactactggagcctggcccagtcactgccaggtggtcctgcaggtgccattc	910
Qy	868	CTCACTACTTACCTGGAGCCTGGCCAGTCACTGCCAGGTGGTCTTCAGAGGCTGCCATTC	927
	911	ctctcaectctgtggctctctcccactgttcaggaccacacagatgggcacaggccaactg	970
Qy	928	CTTCAACCTCTCTGTGGCTCTCTCCCACTTCAGAGCAACACAGATGGGCAAGGCCAAMCTG	987
	971	caagggcccttaacaccacagctggccaaagtgcctactacagaagttgtgggtactacac	1030
Qy	988	CAGAGGCCCTTAACACCAACAGCTGGCCAAAGTGCTACTACAGAAGTTGTGGGTACTACAC	1047
	1031	ctggtcagggcccaactgcagagccctctggaaccacatctgtgcaggtgccacaacactg	1090
Qy	1048	CTGGTCAGGGCCAACTGCAGAGCCCTCTGGAACCAATCTGTGCAGGTGCCAACCACTG	1107
	1091	aagtcataaagcactgcacctgtgcagatgccaaactgcagagagacacaggtatgcacctg	1150
Qy	1108	AAGTCATAAGCACTGCACCTGTGCAGATGCCAACTGCAGAGACACAGGTATGCACCTG	1167
	1151	agaaggtgccagtttcagaggtcatgggtaccacactggcagagatgtcaactccagagg	1210
Qy	1168	AGAAAGTGCCAGTTTCAAGGTCA TGGGTACCACACTGGCAGAGATGCTCAACTCCAGAG	1227
	1211	ctcaggttatgacacctgcagaggtatcaattgtggtgctttctggaaccacagctgcac	1270
Qy	1228	CTACAGGTATGACACCTGCAGAGGTATCAATGTGTGTGCTTCTGCAGACACACAGCTGCAC	1287
	1271	aggtaaacactTACAGAGTGGGTGGACCAACAGCTAGACAGCTACTTATCCCTGAGCGCTG	1347
Qy	1331	aaggtccagatgccagctcaatcatgtctacggaagattaccaggttccctgggcccc	1390
	1348	AGGTCAGATGCCAGCTCAATCATGTCTACGGNAAGTATTACAGGTTCCCTGGGCCCCC	1407
Qy	1391	tgctggatggtacagccaccttaaggtggtgaagagacaagtcacctggatgtgttctc	1450
	1408	TGCTGGATGTGTACAGCCACCTTTAAGCTGTGTTAAGAGACACAGTCCCCCTGGATTTGTCTTC	1467
Qy	1451	tgtatcatgatgtctctttttccgtcacctggacattgtccaggttatataaagatgccg	1510
	1468	TGTATTCGATATGTTCTCTTTTCGCTCAACCTCGACATCTCTCCAGGTTATTGAAAGTGC	1527

Db	1511	agatcctgcaggctgtgcgcgtccggtgaggaggatgcattctgagctgactgtcctgcc	1517
Qy	1528	AGATCTCGAGGCTGTGGCGTCCGCTGCGGATGGAGGGCATGCAATTTGAGCTCACTGTGTCTCGCC	1587
Db	1571	aagcggggctgccaaaggaagcctgcattggagatctcatgccagggtgccagagccctgtg	1630
Qy	1588	AAGCGGGCTGCCAAGGAGCCTGCATCGAGATCTCATGCCAGGTTGCCAGGCCCGCTG	1647
Db	1631	cccagcggtgtgcagcctgtgctaccagccagcctgccagctggctgcacacaga	1690
Qy	1648	CCACAGCGCTGTGCAGCTGTGCTRACCAACCCACCTGCCAGCTGTTCTGCAACGA	1707
Db	1691	tactgaagggtggctcggggacatcctgcctcaatgtctctggctgataccaacagcc	1750
Qy	1708	TACTGAAGGTTGGCTCGGGGACATATCGCTCAATCTGTCTGTGGCTGATACCAACAGCC	1767
Db	1751	tggcagtggtcagcaaccagcttcatcgtcctggctcaagaagcagcccttggcgagtttc	1810
Qy	1768	TGGCAGTGGTCAGCACCCAGCTTATCATGCTGCTGCTCAAGAAGCAGGCGCTTGGCGAGTTTC	1827
Db	1811	cgctgatcgtgggcacatcctgctgggtgtgatgctdggctccttgcatctcgtatatata	1870
Qy	1828	CGCTGATGCTGGGCATCTTCTGGTCTTGTGGCTGTGGTCTCTGCTTGCATCTCTGATATATA	1887
Db	1871	ggcgcagactatgaagcaagacttcctcgtaccagcttgccacatagcagcagtcact	1930
Qy	1888	GGCGCAGACTTATGAGCAAGACTTCTCGGTACCCCACTTGCCACATGACGAGCTCACT	1947
Db	1931	ggctgcgtctaccgcccatctctgctcttgccttgggtgagaaatagccccctcctca	1990
Qy	1948	GGCTGGCTGTACCCCGCATCTTCTGCTTGTCCCATTTGGTGAACACGCCCCCTCCTCA	2007
Db	1991	gtggcgcagagctcgaatcctctcatalgatctgtgatttctcgagtgacagaaa	2050
Qy	2008	GTGGCGCAGCGTCTCAGTACTCTCATATGATGCTGTGATTTTCTGGAGTTTGACAGAA	2067
Db	2051	caactatatccccccagctcttcctctggagactactattaaactgaataaataactcaga	2110
Qy	2068	CAGCTATATTTCCCCCAGCTCTTCCCTGGGAGCTACTATTAACTGAATAAATACTCAGA	2127
Db	2111	gcctg 2115	
Qy	2128	GCCTG 2132	
RESULT 3			
ID	T03760 standard; DNA; 2131 BP.		
AC	T03760;		
DT	25-MAR-1996 (first entry)		
DE	Melanoma-specific immunogen, pMEL17.		
KW	Melanoma; immunogen; epitope; homologue; vaccine; immunotherapy;		
KX	cytotoxic T cell; lymphocyte; HLA-A2; ss.		
OS	Homo sapiens.		
PN	W09522561-A2.		
PD	24-AUG-1995.		
PF	16-FEB-1995; U01991.		
PR	16-FEB-1994; US-197399.		
PS	29-APR-1994; US-234784.		
PA	(U01V1-) UNIV VIRGINIA PATENT FOUND.		
PI	Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CJ;		
DR	WP; 95-30268/39.		
PT	Melanoma-specific immunogen comprises epitope(s) homologous with		
PT	pMEL17 - are highly potent stimulators of HLA-A2+CTL's useful in		



PT adoptive immuno-therapy  
PS Disclosure; Page 19-20; 148pp; English.  
CC A melanoma-specific immunogen homologous with pMel-17 (t03760) comprises  
CC one or more CTL (cytotoxic T lymphocyte) epitopes from the group  
CC R82098-R82194 capable of eliciting a CTL response. The epitopes R82098-  
CC R82108 are of particular interest. The immunogen can be used for  
CC partial protection in mammals against melanoma peptides which are  
CC homologous with pMel-17 are highly potent stimulators of HLA-A2+  
CC CTLs in several cell lines and can be used in immunotherapy or  
CC incorporated into immunogenic conjugates as vaccines. 491 T;  
SQ Sequence 2131 BP; 474 A; 589 C; 577 G;

Query Match 90.7%; Score 1970; DB 17; Length 2131;  
Best Local Similarity 98.7%; Pred. No. 0.00e+00;  
Matches 2103; Conservative 0; Mismatches 7; Indels 21; Gaps 1;

Db	1	ggaagaacacattgatctggtgctaaaaagatgccttcttcatttggctgtagtggtg 60
Qy	28	GGAGACACATGATCTGGTCTTAAAGATGCCCTTCATTTGGCTGTGATAGGTG 87
Db	61	ctttgctggctgtgggggtacaaaagtaccagaaccaggactggttggctgtctcaa 120
Qy	88	CTTTCTGGCTGTGGGGGCTACAAAAGTATCCAGAAACAGAGACTGGCTTGGTGTCTAA 147
Db	121	ggcaactcaagaacaaacctgaaacaggcagctgtatccagatggacagaagccaga 180
Qy	148	GCCAACTCAACAAAGCCGTGGAACAGGAGCTGTATCCAGATGGACAGAGCCGAGA 207
Db	181	gaattgactgctggagagtggtccaagtgtccctcaaggtcagtaatgatggctcaac 240
Qy	208	GACITGACTCTGGAGAGTGGTCAAGTGTCCCTCAAGTCAATATGATGGCTACAC 267
Db	241	tgattggtgcaaatgcctcctctctattgcttgaacttccctggagccaaaagtat 300
Qy	268	TGATTTGGTCAAAATGCCCTCTCTATTATGGCTTGAATTCCTCGGAAGCCAAAAGTAT 327
Db	301	tgccagatggcaggttatctgggtcaaataccatcatcaatggagccaggtgtggg 360
Qy	328	TCCAGATGGCAGGTATCTGGCTACAAATACCATCATATGCGAGCCAGGTGTGG 387
Db	361	gagacagccagtgatccccaggaactgacgatgcttccctgctgattggtgac 420
Qy	388	GAGACAGCCAGTGTATCCCGAGAAACTGACGATGCCCTGCATCTTCCCTGATGGTGAC 447
Db	421	cttgcccatctggtcttggctcagaagagaagcttggttatgctcggaaacctggg 480
Qy	448	CTTGCCCATCTGGCTCTGGTCTCAGAGAGAAGCTTTGTTATGCTCGAAGACCTGG 507
Db	481	gcaatactggcaagtcttagggggccagtgctgggtgagcattggagcagcgagg 540
Qy	508	GCCAACTAGGCAATTTCTAGGGGGCCAGTGTCTGGGCTGAGCATTTGGGACAGGCGG 567
Db	541	caatgctgggcacacacacacattggaagtgactgtctaccatcgccggggatccggagct 600
Qy	568	CAATGCTGGGCGACACACCATGGAAGTGAATCTGTCTACCATCGCCGGGGATCCGGAGCT 627
Db	601	atgtgcctcttgcctcattccagctcagccttcacattactgacccaggtgcctttctcg 660
Qy	628	ATGTGCCTCTTGTCTATTCAGCTCAGCTTTCACATTACTACCAGGTGCTTTTCTCG 687
Db	661	tgagcgtgtcccagttgcccggccttggatogagggaacaagcacttcttgagaatcagc 720
Qy	688	TGAGCGGTCTCCAGTTGGGGGCTTGGATGGAGGGGAACAGCACTTCTCTGAGAAATCAGC 747

Db	721	ctctgaaccttgccctccagctccaatgaacctagtgctctatctggctgaagctgaacct 780
Qy	748	CTCTGACCTTTGGCTCCAGCTCCATGACCCAGTGGCTATCTGGCTGAAGCTGACCTCT 807
Db	781	ctcacacctgggaacttggagacagtagtgaacctgatctctogggacctgtggtca 840
Qy	808	CCTACACCTGGGACTTTGGAGACAGTAGTGGAACTCTGATCTCTGGGCACTTGTGGTCA 867
Db	841	ctcaacttaacctggagcctggccagtcgaactgccaggtggctcctgcaggtgcccattc 900
Qy	868	CTACTACTTACCTTGGAGCCTGGCCAGCTACTCTCCAGGTGGTCTGACAGCTGCCATTC 927
Db	901	ctctcacctcctgtggtcctctcccagttccaggtccaccacagatgggcacaggccaaqtg 960
Qy	928	CTCTCAGCTCTGTGGCTCTCCAGTTCAGGCAACACAGATGGGCACAGGCCCAACTG 987
Db	961	cagaggcccttaaacaccacagctggccaaagtgcctactacaagaagtgtggtaactaac 1020
Qy	988	CAGAGGCCCTTAACACCACAGCTGGCCAAAGTGCCTACTACAGAAAGTTGTGGGTACTACAC 1047
Db	1021	ctggtcaggcgccaaactgcagagcctctggaaccacatctgtgcaagtgccaacctg 1080
Qy	1048	CTGGTCAGGGGCCACTGCGAGAGCCCTCTGGAAACACACATCTGTGCAGGTGCCAACCACTG 1107
Db	1081	aagtataagcactgcacctgtgcagatgccaactgcagagacacaggtatgacacctg 1140
Qy	1108	AGTCATTAAGCACTGCACTGTGCGAGTGGCAACTGCGAGACAGACACAGGTATGACACCTG 1167
Db	1141	agaaggtgccaagtttcagaggtcactgggtaccacactggcagagatgtcaactccagagg 1200
Qy	1168	AGAAGGTGCCAGTTTCAGAGTCTGGGTACACACTGTCGACAGATGTCAACTCCGAGGG 1227
Db	1201	ctcacagtatgacacctgcagaggtatcaattgtgggtgctttctggaaaccaagctgcaac 1260
Qy	1228	CTACAGGTATGACACTGCGAGGTATCAATTTGTGTGCTTCTGCAACACACAGCTGCAC 1287
Db	1261	aggttaacactacagatgggtggagaccacagctagagagctacctatccctgagcctg 1320
Qy	1288	AGCTTAACACTACAGATGGGTGGAGACACACCTTAGAGAGCTACCTATCCCTGAGCCTG 1347
Db	1321	aaggtccagatgccaagctcaatcatgtctacggaaagtattacaggttccctgggcccc 1380
Qy	1348	AAAGTCCAGATGCCAGCTCAATCATGTCTACGGAAAGTATTACAGGTTCCCTGGGCCCCC 1407
Db	1381	tgctggatgtaacgcaaccttaaggtggtgaagagacaagtcacctggattgtgttc 1440
Qy	1408	TGCTGGATGTTACAGCCACTTAAGGCTGGTGAAGACAGAAGTCCGCCCTGCAATGTCTTC 1467
Db	1441	tgctgatatggttctcttccgtcaacctggacatgtccaggggtattgaaagtgcg 1500
Qy	1468	TGTAATCGATATGTTCTTTCCGTTCACCTCGACATTTGTCAGGGGTATTGAAAGTGGCG 1527
Db	1501	agatcctgcaggtgtgcccgtccgggtgaggggagtcatttgagctgactgtgctctgcc 1560
Qy	1528	AGATCCTGCGAGGCTGTGCCGTCCGGTGAGGGGATGCAATTTGAGCTGACTGTGTCTGCC 1587
Db	1561	aagcggggtgcccagaagacctgcatggagatctcatcgccaggtgcccagccccctg 1620
Qy	1588	AAAGCGGGCTGCCAAAGGAAGCCGTGCATGCGAGATCTCATGCCAGGGTCCCGACCCCTG 1647
Db	1621	cccaggggtgtgcccagctgtgctacctaccagccagcctgccagctggttctgcaccaga 1680
Qy	1648	CCGAGCGGCTGTGCCAGCTGTGCTTACTACCCAGCCAGCCTGCCAGCTGGTTCTGCAACGA 1707



RESULT	5	
ID	Q10572	standard; DNA; 1047 BP.
AC	Q10572;	
DT	09-APR-1991	(first entry)
DE	Human Natriuretic Peptide Receptor B.	
KW	NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;	
KW	hyperaldosteronism; glaucoma; guanylyl cyclase.	
OS	Homo sapiens.	
FH	Key	Location/Qualifiers
FT	Peptide	1..22
FT	/label= signal sequence	
FT	Protein	12
FT	/label= mature NPRB	23..455
FT	Domain	
FT	/label= extracellular domain	
FT	/note= "binds natriuretic peptides A, B and C"	
FT	Domain	456..456
FT	/label= transmembrane domain	
FT	Domain	479..1047
FT	/label= cytoplasmic domain	
FT	/note= "GC and protien kinase activity"	
FT	Modified-site	24..26
FT	/label= N-glycos site	
FT	Modified-site	35..37
FT	/label= N-glycos site	
FT	Modified-site	161..163
FT	/label= N-glycos site	

FT	Modified -site 195..197
FT	/label= N-glycos site
FT	Modified -site 244..246
FT	/label= N-glycos site
FT	Modified -site 277..279
FT	/label= N-glycos site
FT	Modified -site 349..351
FT	/label= N-glycos site
FT	Modified -site 600..602
PN	/label= N-glycos_site
PD	#09100292-A.
PF	10-JAN-1991.
PP	22-JUN-1990; U03586.
PR	23-JUN-1989; US-370673.
PA	(GETH) GENENTECH INC.
PI	Chang M, Goeddel D, Lowe D;
PJ	WPI; 91-036711/05.
DR	N-PSDB; Q10324.
PT	Natriuretic protein receptor B
PT	Natriuretic protein receptor B - for diagnosis and treatment of
PT	kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS	Claim 3; Fig 1; 49pp; English.
CC	The sequence was derived from the DNA encoding natriuretic peptide
CC	receptor B. NPRB, having guanyl cyclase (GC) activity and protein
CC	kinase activity. The DNA can be inserted into expression vectors
CC	for the prodn. of the protein, opt. after being mutated to produce
CC	NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=
CC	114,952). The protein (or variants) can be used in treatment of
CC	natriuretic peptide disorders, and also to isolate peptides using
CC	affinity chromatography. Antibodies with affinity for NPRB can
CC	also be prepd.
SQ	Sequence 1047 BP; 87 A; 15 C; 83 G; 51 T;
Query Match 4.0%; Score 86; DB 2; Length 1047;	
Best Local Similarity 9.5%; Pred. No. 3.04e-28;	
Matches 92; Conservative 279; Mismatches 581; Indels 15; Gaps 15;	
Db	42 wnrvgnavanvgrannvdrrrvssnnngacsnynnaansavdknyhdndnnngc 101
Cp	1498 AGGTCACGGAAAGGACATATCGATACAGAACAATCCAGGGGACTTCTCTTC 1439
Db	102 vynaasvarnaeshwrnnntagavasgne-akndhytnrvrtgn-sankngnnvnvtbhg 159
Cp	1438 ACCAGCCTTAAGTGCGCTGTACCATTCCAGCAGGGGCCAGGGAACCTGTAACTTTCC 1379
Db	160 mwttaaanvnyndartddrhnytmgvnnannsnsvnhnvayrningmnmnmra 219
Cp	1378 GTACACATCATTTAGCTGGCATCTGCACCTTCAGGCTCAGGATAGTAGCTCTAGCT 1319
Db	220 ngtrvnycegnmmhhnnmnnaurnrntngdyvnyndvngnsrnaagratgrnwdnrt 279
Cp	1318 GTGETCTCACCCACTCTGTAGTTGTACTCTGTGACCTGTGTTTCAGAAAGCAACACA 1259
Db	280 rmaanaantvntvyrnnnnnyyrrnnmrnrnrnrndingvngnsnmnnnagcnvd 339
Cp	1258 ATTGANTACTCTGCAGGCTGTATACTTGTAGCTGTAGCTCTCGAAGTTCAGCTCTG 1199
Db	340 gnnny-anvrnnntnnngtrndngprnvmkmgrgyrghvqnvvmkdnn-drn-tdnvwa 396
Cp	1198 GTACCATGAGCCTCTGAACCTGGCACCTTCTCAGGTGTACATCTGTGCTCTGCAGTT 1139
Db	397 mgdrdsdgndnaahysganknwttgrnnnwkykanns dnnncandndndscdktnnstn 456
Cp	1138 GGCACTGCAGAGCTGCAGCTGTATGACTTTCAG-TGGTTGGCAGCTGCACAGATGCTGT 1080



FI CDS  
17.1.1204

Q Sequence 1/10 BP; 534 A; 331 C; 364 G; 415 T;

SQ Sequence 1710 BP; 534 A; 337 C; 364 G; 475 T;

```

Query Match      2.0%; Score 44; DB 8; Length 1710;
Best Local Similarity 78.2%; Pred. No. 6.69e-08;
Matches 61; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Db    1621 ggagattaatatatttataaaaaaaaaaaaaaaaaaaaaaa 1680
      ||||| || | | | |||||| | | | ||||||| |||||||
Qy    2095 GGAGACTACTATTAACTGAATAAATACCTCAGCGCTGAAAAAATAAAAAAAAAAAAAA 2154

Db    1681 aaaaaaaaaaaaaaaaaa 1698
      ||||||| ||||||| |||||
Qy    2155 AAAAAAAAAAAAAAAAAA 2172

RESULT 11
ID Q88760 standard; DNA; 1813 BP.
AC Q88760;
DT 20-JAN-1996 (first entry)
DE Human ubiquitous nuclear receptor protein DNA.
KW Ubiquitous nuclear receptor; drug design; disease diagnosis;
KW therapy; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1..1727
   /tag= a
   PN W09513373-A1.
   PD 18-MAY-1995.
   PF 08-NOV-1994; U12883.
   PR 10-NOV-1993; US-152003.
   PP (ARCH-) ARCH DEV CORP.
   PI Liao S.; Song G;
   DP WPI; 95-194092/25.
   DR P-PDSB; R14738.
   New ubiquitous nuclear receptor - used to develop prods. for use in
   diagnostic, drug design and therapeutic applications
   Claim 6; Page 148; 195pp; English.
   This genomic DNA isolated from human prostate cancer PC-3 cells may
   be expressed recombinantly to produce a UNRP which interacts with
   the response elements and network of receptors in the thyroid
   hormone/retinoic acid receptor subfamily. The products can be used
   in detection, diagnosis and in screening assays for substances which
   interact with UNRPs for use in diagnosis, drug design and
   therapeutic applications.
   Sequence 1813 BP; 410 A; 573 C; 529 G; 301 T;

SQ
Query Match      2.0%; Score 43; DB 15; Length 1813;
Best Local Similarity 86.4%; Pred. No. 1.88e-07;
Matches 51; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Db    1728 aataaaaacataaacagaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1786
      ||||| || | | | |||||| | | | ||||||| |||||||
Qy    2114 ATAAATACCTCAGCGCTGAAAAAATAAAAAAAAAAAAAAAAAAAAAA 2172

RESULT 12
ID Q51746 standard; cDNA; 91 BP.
AC Q51746;
DT 31-MAY-1994 (first entry)
DE Oligonucleotide probe MK14-A
KW oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW ss.
OS Synthetic.
PN EP-571911-A.
PD 01-DEC-1993.
```

DE	Human superoxide dismutase gene.	
DS	Superoxide dismutase; SOD; transformation; production; ss.	
KW	Homo sapiens.	
FH	Key	Location/Qualifiers
FT	CDS	68..529
FT	/*tag= a	
FT	/product= superoxide dismutase	
FT	polya_signal	625..699
FT	/*tag= b	
FT	J06046860-A.	
PN	PN	
PD	22-FEB-1994.	
PF	25-SEP-1992; 279193.	
PP	25-SEP-1992; JP-279193.	
PR	(NIPK ) NIPPON KAYAKU KK.	
PA	WPI; 94-097023/12.	
DR	P-PSDB; R47062.	
DP	DNA encoding human superoxide dismutase (SOD) and transformed	
PT	microorganism - useful for prodn. of SOD	
PT	Claim 1; Page 7; 10pp; Japanese	
PP	The sequence encodes a superoxide dismutase. The gene may be used to	
CC	produce superoxide dismutase in increased yields by transforming it	
CC	into E.coli.	
CC	Sequence	699 BP; 242 A; 127 C; 181 G; 149 T;

Query Match 1.9%; Score 41; DB 10; Length 699;  
Best Local Similarity 86.0%; Pred. No. 1.46e-06;  
Matches 49; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Db 609 taacactgtatcttaaaaaaaataaaaaaaaataaaaaaaa 665  
 ||||| ||| - | ||||| ||||| ||||| ||||| ||||| |||||  
 Qv 2116 TAAATACTCAGCGCTGAAAAAATAAAAAAAAAAAAAAAAAA 2172

RESULT	15
ID	CQ73729 standard; cDNA; 756 BP.
AC	Q73729;
DT	26-APR-1995 (first entry)
DE	Rape abscission/dehiscence protein pSACS1 cDNA.
KW	Rape; pSACS1; abscission; dehiscence; transgenic plant;
KW	crop improvement; ss.
OS	Brassica napus.
FH	Key
FT	Location/Qualifiers
FT	15..458
FT	/*tag= a
FT	polya signal 633..638
FT	/*tag= b
FT	/note= "putative polyA site"
FW	M09423043-A.
PDP	13-OCT-1994.
PPD	31-MAR-1994; G00689.
PE	31-MAR-1994; GB-006726.
PR	(NICK-) NICKERSON BIOGEN LTD.
PI	Coupe SA, Isaac PG, Roberts JA;
PL	MP1; 94-333203/41.
DR	P-P5DB; R60799.
DPT	Recombinant and isolated nucleic acids - encode enzymes and
TPT	Disclosures involved in plant abscission or dehiscence.
SPTS	Disclosure; Fig 3; 39pp; English.
CC	The cDNA encodes a rape (cv. Rafal) protein involved in abscission
CC	or dehiscence, and can be used to manipulate plant abscission and
CC	reduce/prevent pod dehiscence, especially to regulate abscission of
CC	pollen from anthers, to generate male sterile plants and for hybrid
CC	seed production.
SC	Sequence 756 BP; 240 A; 175 C; 118 G; 223 T;

Query Match	1.9%;	Score 41;	DB 12;	Length 756;
Best Local Similarity	97.7%;	Pred. No. 1.46e-06;		
Matches	42;	Conservative	0;	Mismatches 1; Indels 0; Gaps 0;

Search completed: Tue Aug 27 09:12:38 1996  
Job time : 188 secs.



\*\*\*\*\*

WAVELENGTH (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.

Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.

Distribution rights by IntelliGenetics, Inc.

MPSrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Aug 27 09:13:00 1996; MasPar time 847.32 Seconds  
921.604 Million cell updates/sec

Tabular output not generated.

Title: &gt;US-08-231-565A-26

Description: (1-2172) from US08231565A.seq

Perfect Score: 2172

N.A. Sequence: 1 GTCGACGGCCATTACCAATC.....AAAAAAAAAAAAAAAAAAAA 2172

Comp: CACTGCGCGGTATGCTAG.....TTTTTTTTTTTTTTTTTT

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 514334 segs, 179763086 bases x 2

Post-processing: Minimum Match 04

Listing first 45 summaries

Database:

EST-STS  
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8  
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14  
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20  
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26  
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32  
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38  
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44  
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50  
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56  
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62  
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68  
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74  
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80  
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86  
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:STS2  
93:STS3 94:STS4 95:STS5 96:STS6 97:STS7  
EST-STS-TWO  
98:EST91 99:EST92 100:EST93 101:EST94 102:EST95  
103:EST96 104:EST97 105:EST98 106:EST99 107:EST100  
108:EST101 109:EST110 110:EST111 111:EST112 112:EST113  
113:EST114 114:EST115 115:EST116 116:EST117 117:EST118  
118:EST119 119:EST120 120:EST121 121:EST122 122:EST123  
123:EST124 124:EST125 125:EST126 126:EST127  
127:EST128 128:EST129 129:EST130 130:EST131 131:EST132

Database:

98:EST91 99:EST92 100:EST93 101:EST94 102:EST95  
103:EST96 104:EST97 105:EST98 106:EST99 107:EST100  
108:EST101 109:EST110 110:EST111 111:EST112 112:EST113  
113:EST114 114:EST115 115:EST116 116:EST117 117:EST118  
118:EST119 119:EST120 120:EST121 121:EST122 122:EST123  
123:EST124 124:EST125 125:EST126 126:EST127  
127:EST128 128:EST129 129:EST130 130:EST131 131:EST132

132:enEST7 133:enEST8 134:enEST9 135:enEST10 136:enEST11  
137:enEST12

Statistics: Mean 12.077; Variance 3.893; scale 3.102

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	542	25.0	621	114	N28728	yx67al2.r1 Homo sapie 0.00e+00
2	542	25.0	621	136	HS728262	yx67al2.r1 Homo sapie 0.00e+00
3	496	22.8	638	114	N28736	yx67cl2.r1 Homo sapie 0.00e+00
4	496	22.8	638	136	HS736264	yx67cl2.r1 Homo sapie 0.00e+00
5	426	19.6	546	136	HS746262	yx67fl1.r1 Homo sapie 0.00e+00
6	426	19.6	546	114	N28746	yx67fl1.r1 Homo sapie 0.00e+00
c 7	420	19.3	430	130	HS118284	yx67cl2.s1 Homo sapie 0.00e+00
c 8	420	19.3	430	113	N23118	yx67cl2.s1 Homo sapie 0.00e+00
c 9	405	18.6	431	113	N23111	yx67al2.s1 Homo sapie 0.00e+00
c 10	405	18.6	431	129	HS111256	yx67al2.s1 Homo sapie 0.00e+00
c 11	380	17.5	521	118	N40290	yx80c07.r1 Homo sapie 0.00e+00
12	380	17.5	521	131	HS290283	yx80c07.r1 Homo sapie 0.00e+00
13	351	16.2	416	68	R78477	y173d06.r1 Homo sapie 0.00e+00
14	273	12.6	306	83	T29726	EST92123 Homo sapiens 4.76e-298
15	258	11.9	286	46	R01509	ye79a02.r1 Homo sapie 9.86e-279
c 16	54	2.5	58	113	N23128	yx67fl1.s1 Homo sapie 1.47e-28
c 17	54	2.5	58	130	HS128259	yx67fl1.s1 Homo sapie 1.47e-28
18	43	2.0	139	82	T25552	EST00585 Mus musculus 2.58e-17
19	43	2.0	141	119	MUSGS00442	Mouse 3'-directed cd 2.58e-17
20	43	2.0	141	45	MUSGS00442	Mouse 3'-directed cd 2.58e-17
21	43	2.0	319	66	R71994	yj84b09.r1 Homo sapie 2.58e-17
22	43	2.0	336	69	R82874	yj25g09.r1 Homo sapie 2.58e-17
23	43	2.0	409	119	N42987	yy08b05.r1 Homo sapie 2.58e-17
24	41	1.9	151	78	T04695	742 Arabidopsis thali 2.29e-15
25	41	1.9	153	130	HS165282	yy84f10.r1 Homo sapie 2.29e-15
26	42	1.9	207	64	R66174	yi30c05.r1 Homo sapie 2.45e-16
c 27	41	1.9	208	8	HL4237	ym63d10.r1 Homo sapie 2.29e-15
28	42	1.9	219	69	R81427	yj02a05.r1 Homo sapie 2.45e-16
29	42	1.9	260	65	R69938	yi47q01.r1 Homo sapie 2.45e-16
c 30	41	1.9	281	132	HS341251	yw33g02.s1 Homo sapie 2.29e-15
c 31	41	1.9	281	112	N22341	yw33g02.s1 Homo sapie 2.29e-15
c 32	42	1.9	333	5	H06053	yl72b12.s1 Homo sapie 2.45e-16
c 33	41	1.9	336	109	H92687	yt90b07.r1 Homo sapie 2.29e-15
c 34	41	1.9	347	1	ATTS2436	A. thaliana transcrib 2.29e-15
c 35	41	1.9	360	131	HS271255	yw47g05.s1 Homo sapie 2.29e-15
c 36	41	1.9	360	113	N23271	yw47g05.s1 Homo sapie 2.29e-15
37	42	1.9	391	125	G15952	human STS CHLC.UTR 04 2.45e-16
38	41	1.9	417	72	R92529	yq08c01.r1 Homo sapie 2.29e-15
c 39	41	1.9	427	24	H67242	yq65g03.r1 Homo sapie 2.29e-15
c 40	41	1.9	430	59	R49270	yq59b02.s1 Homo sapie 2.29e-15
c 41	42	1.9	444	130	HS135291	yz42c12.s1 Homo sapie 2.45e-16
c 42	41	1.9	468	8	H14523	yl25a09.r1 Homo sapie 2.29e-15
c 43	40	1.8	295	118	N39304	yv25b02.s1 Homo sapie 2.09e-14
c 44	40	1.8	448	56	R37188	yh96a12.r1 Homo sapie 2.09e-14
c 45	40	1.8	485	60	R51745	yg73c12.s1 Homo sapie 2.09e-14

## ALIGNMENTS

RESULT 1



Aug 27 09:14

US-08-231-565A-26.rst

3

LOCUS N28728 621 bp mRNA EST 04-JAN-1996  
DEFINITION Y67a12.r1 Homo sapiens cDNA clone 266782 5' similar to  
SP:PIR:A53668 A53668 glycoprotein gp100 precursor, melanocyte  
lineage - human ;  
ACCESSION N28728  
NID g1146964  
KEYWORDS EST.  
SOURCE human clones-266782 primer=T7 library=Soares melanocyte 2NblHM  
vector=pT7T3D (Pharmacia) with a modified polylinker host=PH10B  
(ampicillin resistant) Reitel=Not I Reite2=Eco RI Male. 1st strand  
cDNA was primed with a Not I - oligo (dT) primer  
[5'-TGTTACCAATCTGAAGTGGAGCGCGCGAGTTTTTTTTTTTTTTT-3'],  
double-stranded cDNA was size selected, ligated to Eco RI adapters  
(Pharmacia), digested with Not I and cloned into the Not I and Eco  
RI sites of a modified pT7T3 vector (Pharmacia). Library  
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal  
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.  
Albino.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;  
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;  
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 621)

## REFERENCE

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and  
Wilson, R.

## TITLE

The WashU-Merck EST Project

## JOURNAL

Unpublished (1995)

## COMMENT

Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
High quality sequence stops: 398  
Source: IMAGE Consortium, L1NL  
This clone is available royalty-free through L1NL; contact the  
IMAGE Consortium (info@image.l1nl.gov) for further information.

NCBI gi: 1146964

## FEATURES

## source

Location/Qualifiers

1..621  
/organism="Homo sapiens"  
/clone="266782"  
/note="human"  
<1..>621

## BASE COUNT

125 a 170 c 173 g 145 t 8 others

## ORIGIN

Query Match 25.0%; Score 542; DB 114; Length 621;  
Best Local Similarity 96.2%; Pred. No. 0.00e+00;  
Matches 604; Conservative 0; Mismatches 15; Indels 9; Gaps 9;  
Db 1 ctgagagctacctccctgagctgaaggtccagatgccagctcaatcatgtctacgg 60  
|||||  
Qy 1321 CTGAGAGCTACCTATCCCTGAGGCTGAAGTCCAGATGCCAGCTCAATCATGTCTACGG 1380  
|||||  
Db 61 aaagtattacaggttcctctggccccctgctggatgtacagccaccttaagctgtgga 120  
|||||

Aug 27 09:14

US-08-231-565A-26.rst

4

Qy 1381 AAGTATTACAGTTCCCTCGGCCCTCTGCTGATGCTAGCAGCCACTTAAGCGTGGTCA 1440  
Db 121 agagacaagtccccctgattgtgtctgtatcatgatatgggtctctttccgtcacctcgg 180  
|||||  
Qy 1441 AGAGACAAGTCCCGCTGGATTGTGTCTGATCGATATGGTCTCTTTTCCGTCACCCCTGG 1500  
Db 181 acattgtccagatnaintgaaagtccagatcctgcaggctgtgcgtncsggtgagggg 240  
|||||  
Qy 1501 ACATTTGTCAGGGTAT-TGAAGTGCAGATCTCCAGGCTGTCCGCTCCGCTGAGGGG 1559  
Db 241 gatcatttgagctgactgtgtctgccaaaggcgggtgcccaaggaaagcctgcattgag 300  
|||||  
Qy 1560 GATGCATTTCAGCTGACTGTGTCTGCCAAGGGGGCTGCCCAAGGAAGCCTGCA TGGAG 1619  
Db 301 atctcatcgccagggtgccagccctgccagcggtgtgc-agcctgtgtaccagc 359  
|||||  
Qy 1620 ATCTCATGCCAGGGTGCAGCCCTCTGCCAGCGGTGCGCAGCGCTGTGCTTACCCAGC 1679  
Db 360 ccagctgccagctggtttctgcaccagatactgaagggtggtcggggacatactgcctc 419  
|||||  
Qy 1680 CCAGGCTGCCAGCTGCTTCTGCACCATATCTCAAGGGTGGCTCGGGGACATATCTGCCTC 1739  
Db 420 aatgtctctggtgatataccaaagcctggtgcagtggtcagcaaaccaagcttcatcgct 479  
|||||  
Qy 1740 AATGTGTCTCTGGCTGTATACCAAGCCTGGCAGTGGTTCAGCACCAGCTTATCATGCTC 1799  
Db 480 ggtcaagaagcggcttggnaaggttcctgctcagtgatccgtgggcantttggtgtgtgat 539  
|||||  
Qy 1800 GGTCAAGACAGCGCTTGGGCGAGGTTCCTGCTATC-CTGGGCACTTCTGCTGTGTGAT 1858  
Db 540 ggctgtggcctttcatctctgntatataggc-ca-acttatgaagcaa-actt-tnogt 595  
|||||  
Qy 1859 GGCTGTGCTCTTGCATCTCTGATATATAGCGCAGACTTATGAAGCAAGACTTCTCGGT 1918  
Db 596 accc-agttgc-acatagcagcagcnaac 621  
|||||  
Qy 1919 ACCCCAGTTGCCACATAGCAGCAGTCAC 1946

## RESULT

2

ID HS728262 standard; RNA; EST; 621 BP.  
AC N28728;  
DT 05-JAN-1996 (Rel. 46, Created)  
DT 05-JAN-1996 (Rel. 46, Last updated, Version 1)  
DE y67a12.r1 Homo sapiens cDNA clone 266782 5' similar to  
DE SP:PIR:A53668 A53668 glycoprotein gp100 precursor, melanocyte  
DE lineage - human ;  
KW EST.  
OS Homo sapiens (human)  
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;  
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.  
RN [1]  
RP 1-621  
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,  
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,  
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,  
RA Trevaskis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;  
RT "The WashU-Merck EST Project";  
RL Unpublished.  
CC Contact: Wilson RK WashU-Merck EST Project Washington University  
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,  
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:  
CC est@watson.wustl.edu High quality sequence stops: 398 Source: IMAGE  
CC Consortium, L1NL This clone is available royalty-free through L1NL

CC	; contact the IMAGE Consortium (info@image.llnl.gov) for further
CC	information. NCBI gi: 1146964
FH	Key Location/Qualifiers

lineage - human ;  
ACCESSION N28736  
NID g1146972  
KEYWORDS EST.

```

QY 1441 ACAGACAAGTCCCGCTGGATTGTTCTGTATCGATATGGTTCTTTTCCTGCACCCCTGG 1500
|||||
Db 181 acattgtccaggttatttgaagtcgcagatcctgcaggtgtgctgtnccgtgagggg 240
|||||
QY 1501 ACATTTGTCAGGGTATT-GAAATGTCGAGATCTTGAGGGTGTGCGCTGCGGTGAGGGG 1559
|||||
Db 241 qatcattgaactgaactgtgctgccaagcgagctgccaaggaagcctgcatgag 300
|||||
QY 1560 GATGCATTTGAGCTGACTGTGTCTGCTGCAAGGGGGCTGCCAAGAACCTGCATGGAG 1619
|||||
Db 301 atctcatgcccaggtgcccagccccctgcccagcggtgtgc-agcctgtgtctaccagc 359
|||||
QY 1620 ATCTCATGCGCAGGTGCCAGGCCCTGCCAGGGCTGTGCCAGGCTGTCTTACCAGC 1679
|||||
Db 360 ccagctgcagctggttctgaccagatactgaaggtggtgctcggggaacataetgct 419
|||||
QY 1680 CCAGCCTGCCAGCTGGTCTGCACACATCTGAAGGGTGGCTCGGGGA-CATACCTGCT 1738
|||||
Db 420 caatgtctctggtgataacaacagcctgacgtgtgcagccagcttatcatgc 479
|||||
QY 1739 CAATGTCTGTGGTGTATACCAA-CAGCCTGGCAGTGGTGCAGCCAGCTTATCATGC 1797
|||||
Db 480 ctggtcaagaacagcctgnaaggtccgtgctgaggggagctctgctggtgta 539
|||||
QY 1798 CTGGTCAAGACGGCTTGGCGAGTTCCGCTGATCGTGGG-CATCTTCTGCTGTTG 1856
|||||
Db 540 atgctgtgg-ccttgnacccttataaaatagg 571
|||||
QY 1857 ATGGCTGTGCTTCGATCTCTGATATAGG 1889
|||||

```

## RESULT 4

```

ID HS736264 standard; RNA; EST; 638 BP.
AC N28736;
DT 05-JAN-1996 (Rel. 46, Created)
DE yx67c12.rl Homo sapiens cDNA clone 266806 5' similar to
DE SP:PIR:A53668 A53668 glycoprotein gp100 precursor, melanocyte
DE lineage - human ;.
KW EST.
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
RN [1]
RP 1-638
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA Trevaskis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;
RT "The WashU-Merck EST Project";
RL Unpublished.
CC Contact: Wilson RK WashU-Merck EST Project Washington University
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC est@wustl.wustl.edu High quality sequence stops: 368 Source: IMAGE
CC Consortium, LIND. This clone is available royalty-free through LIND
CC ; contact the IMAGE Consortium (info@image.lind.gov) for further
CC information. NCBI gi: 1146972
FH Key Location/Qualifiers
FT source
FT 1..638
FT /organism="Homo sapiens"
FT /clone="266806"

```

```

FT mRNA /note="human"
SQ Sequence 638 BP; 133 A; 169 C; 173 G; 153 T; 10 other;

Query Match 22.8%; Score 496; DB 136; Length 638;
Best Local Similarity 95.5%; Pred. No. 0.00e+00;
Matches 547; Conservative 0; Mismatches 20; Indels 6; Gaps 6;

Db 1 ctagagagctactatccctgagcctgaaggtccagatgccagctcaatcatgtctacgg 60
|||||
QY 1321 CTAGAGAGCTACTATCCCTGAGCCTGAAGTCCAGATGCCAGCTCAATCATGTCTACGG 1380
|||||
Db 61 aaagtattacaggttccctggccccctgctggtggtacagccaccttaaggtgtgga 120
|||||
QY 1381 AAGATATTACAGGTTCCTGGGCCCCCTGCTGGATGTGTACAGCCACCTTAAGGCTGTGA 1440
|||||
Db 121 agagacaagtcctccctgattgtgtctatcgatatggttctcttccgtcacctgg 180
|||||
QY 1441 AGAGACAAGTCCCTGGATTGTGTCTGTATCGATATGGTCTTTTCCGTACCCCTGG 1500
|||||
Db 181 acattgtccaggttatttgaagtcgcagatcctgcaggtgtgctgncggtgagggg 240
|||||
QY 1501 ACATTTGTCAGGGTATT-GAAATGTCGAGATCTTCCAGGCTGTGCCGTGAGGGG 1559
|||||
Db 241 gatgcattgagctgaactgtgctctgccaagcggtggtgccaaggaagcctgcatgag 300
|||||
QY 1560 GATGCATTTGAGCTGACTGTGTCTGCTGCCAAGGGGGCTGCCAAGGAAGCCTGCATGGAG 1619
|||||
Db 301 atctcatgcccaggtgcccagccccctgcccagcggtgtgc-agcctgtgtctaccagc 359
|||||
QY 1620 ATCTCATGCGCAGGTGCCAGGCCCTGCCAGCGCTGTGCCAGGCTGTGCTTACCAGC 1679
|||||
Db 360 ccagctgcagcagctggttctgaccagatactgaaggtggtgctcggggaacataetgct 419
|||||
QY 1680 CCAGCCTGCCAGCTGGTTCGTGCACACATCTGAAGGGTGGCTCGGGGA-CATACCTGCT 1738
|||||
Db 420 caatgtctctggtgataacaacagcctgacgtgtgcagcctggttcagccagcttatcatgc 479
|||||
QY 1739 CAATGTCTGTGGTGTATACCAA-CAGCCTGGCAGTGGTGCAGCCAGCTTATCATGC 1797
|||||
Db 480 ctggtcaagaacagcctgnaaggttccgtgctgaggggagctctgctggtgta 539
|||||
QY 1798 CTGGTCAAGACGGCTTGGCGAGTTCCGCTGATCGTGGG-CATCTTCTGCTGTTG 1856
|||||
Db 540 atgctgtgg-ccttgnacccttataaaatagg 571
|||||
QY 1857 ATGGCTGTGCTTCGATCTCTGATATAGG 1889
|||||

```

## RESULT 5

```

ID HS746262 standard; RNA; EST; 546 BP.
AC N28746;
DT 05-JAN-1996 (Rel. 46, Created)
DE yx67f11.rl Homo sapiens cDNA clone 266829 5' similar to
DE SP:PIR:A53668 A53668 glycoprotein gp100 precursor, melanocyte
DE lineage - human ;.
KW EST.
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
RN [1]
RP 1-546
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,

```

RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,  
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,  
RA Trevaskis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;  
RT "The WashU-Merck EST Project";  
RL Unpublished.

CC Contact: Willson RK WashU-Merck EST Project Washington University  
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,  
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:

CC est@watson.wustl.edu High quality sequence stops: 377 Source: IMAGE  
CC Consortium, LNL This clone is available royalty-free through LNL  
CC ; contact the IMAGE Consortium (info@image.llnl.gov) for further  
CC information. NCBI gi: 1146982

FF Key Location/Qualifiers

FN

FT source

FT 1..546

FT /organism="Homo sapiens"

FT /clone="266829"

FT /note="human"

FT mRNA

FT <1..>546

FT

SQ Sequence 546 bp; 114 A; 145 C; 158 G; 120 T; 9 other;

Query Match

Best Local Similarity 94.8%; Pred. No. 0.00e+00;

Matches 507; Conservative 0; Mismatches 16; Indels 12; Gaps 11;

Db 2 catntacagagtgggtggagacacagctagagactacattatccctgagctgaagtc 61

|| |||||

Qy 1294 CAACTACAGAGTGGGTGGAGACACAGCTAGAGCTACCTATCCCTGAGCTGAAGGTC 1353

Db 62 cagatgccagctcaatcatgtctacggaaagtattacaggttccctgggcccctgtcgg 121

|||||

Qy 1354 CAGATGCCAGCTCAATCATCTCTACGAAAGTATTACAGTTCCCTGGGGCCCTGCTGG 1413

Db 122 atgtacagaccacttaagcgtggtgaagagacaagtcaccttgattgtctgtatc 181

|||||

Qy 1414 ATGCTACAGCCACCTTAAGCCTGTGTAAGACAGCAAGTCCCTGTGATTTCTGTATC 1473

Db 182 gatatgttcttttcgtcaccttgacattgtccaggttnatntgaaagtgcagatc 241

|||||

Qy 1474 GATATGTTCTTTTCCTGTCACCTGGACATTTGTCAGGGTAT-TGAAGTGGCAGATC 1532

Db 242 ctgcaggtgtcgtcccggtgaggggatgcatcttgagctgactgtcctgccaaaggc 301

|||||

Qy 1533 CTGCAGGCTGTCCCTCCGCTGAGGGGATGCATTTGAGCTGACTGTCTCTGCCAAGGC 1592

Db 302 gggctgccaaagagcgtgcattgagatcatcagcagggtgccagccctgccag 361

|||||

Qy 1593 GGGCTGCCAAGAGAGCTGCATGAGATCTCATGCGCAGGGTGCAGCCCTGCCAG 1652

Db 362 cggctgtgc-acctgtgtaccagccagcgtgccagctggttctgccaccagatc 420

|||||

Qy 1653 CGGCTGTGCAGCCTGTGCTTACCAGCCAGCCTGCCAGCTGTTCTGCACAG-ATACT 1711

Db 421 gaagggtggctcggggaaacataactggctcaatgtctctctgtgctggatanaacagc 480

|||||

Qy 1712 GAAAGG-TGGCTCGGGG--ACATA-CTGCCTCAATGTCTCTGTGGTG-ATACCACAGC 1766

Db 481 ctggcagtggttnaagcaacccactttatcatgtnctgggncagaagaacaggnctt 535

|||||

Qy 1767 CTGGCACT-GGTCA-GCA-CCAGCTTATCATGCTCTGG-TCAAGAGCAGGCGCTT 1817

RESULT 6

LOCUS

N28746

546 bp

mRNA

EST

04-JUN-1996

DEFINITION

yx67f11.r1 Homo sapiens cDNA clone 266829 5' similar to

SP:PIR:A53668 A53668 glycoprotein gp100 precursor, melanocyte

lineage - human ;

ACCESSION

N28746

NID

g1146982

KEYWORDS

EST.

SOURCE

human clone=266829 primer=T7 library=Soares melanocyte 2NbhM

vector=pfT7D (Pharmacia) with a modified polylinker host=DH10B

(ampicillin resistant) Reitel=Not I Reite2=Eco RI Male. 1st strand

cDNA was primed with a Not I - oligo(dT) primer

[5'-TGTTACCAATCTGAAGTGGAGCGCGAGTTTTTTTTTTTTTTT-3'],

double-stranded cDNA was size selected, ligated to Eco RI adapters

(Pharmacia), digested with Not I and cloned into the Not I and Eco

RI sites of a modified pT73 vector (Pharmacia). Library

constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal

foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.

Albino.

ORGANISM

Homo sapiens

Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;

Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;

Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;

Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 546)

AUTHORS

Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,

Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,

Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,

Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and

Wilson,R.

TITLE

The WashU-Merck EST Project

UNPUBLISHED (1995)

CONTACT: Willson RK

WashU-Merck EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

High quality sequence stops: 377

Source: IMAGE Consortium, LNL

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 1146982

Location/Qualifiers

1..546

/organism="Homo sapiens"

/clone="266829"

/note="human"

<1..>546

BASE COUNT

114 a 145 c 158 g 120 t 9 others

ORIGIN

Query Match

Best Local Similarity 94.8%; Pred. No. 0.00e+00;

Matches 507; Conservative 0; Mismatches 16; Indels 12; Gaps 11;

Db 2 catntacagagtgggtggagacacagctagagactacattatccctgagctgaagtc 61

|| |||||

Qy 1294 CAACTACAGAGTGGGTGGAGACACAGCTAGAGCTACCTATCCCTGAGCTGAAGGTC 1353

Db 62 cagatgccagctcaatcatgtctacggaaagtattacaggttccctgggcccctgtcgg 121

|||||

Qy 1354 CAGATGCCAGCTCAATCATCTCTACGAAAGTATTACAGGTTCCCTGGGGCCCTGCTGG 1413

Db 122 atggtacagccacccctaaagctggtgaagagacagcccccctggattgtgtctgtatc 181  
|||||  
Qy 1414 ATGTGACAGCCACCTTAAAGCTGGTGAAGAGCAAGTCCCCCTGGATTGTGTGTATC 1473  
|||||  
Db 182 gataatgttctcttccgtcacccctggaacatgtccaggtgnaatgaaagtgcagagac 241  
|||||  
Qy 1474 GATATGTTCTTTTCCGCTCACCTGGACATTGTCCAGGTAT-TCAAAGTGGCCGAGTC 1532  
|||||  
Db 242 ctgcagagctgctcccggtgaggggagatgcaatttgactgtctcctccaaggc 301  
|||||  
Qy 1533 CTGCAGGCTGTCCCTCGCTGGGGGATGCAATTGAGCTGACTGTCTCTGCCAAGGC 1592  
|||||  
Db 302 gggctgcccaaggagcctgcatggagatctcatcgccaggtgcccagccccctgccag 361  
|||||  
Qy 1593 GGGCTGCCCAAGGAGGCTGCATGGAGATCTCATGCCAGGGTGGCCGCCCTGCCAG 1652  
|||||  
Db 362 cggctgtgc-agcctgtctaccagcccagcctgccagctggtctcaccagagatc 420  
|||||  
Qy 1653 CGGCTGGCAGCCTGTCTACCCAGCCAGCCTGCCAGCTGGTCTGCACGAG-ATACT 1711  
|||||  
Db 421 gaaggggtgctcggggaacataactggctcaatgtctcgtgctgctgctgatancaacagc 480  
|||||  
Qy 1712 GAAAGG-TGGCTCGGGG--ACATA-CTGCCCTCAATGTGTCTCTGGCTG-ATACCAACAGC 1766  
|||||  
Db 481 ctggcaggtgtnaagcaaccccaactttatcatcgtcgtgggncagaagaacagcgtt 535  
|||||  
Qy 1767 CTGGCAGT-GGTCA-GCA-CCCAGCTTATCATGCTCTGG-TCAAGAAGCAGCGCTT 1817  
|||||

## RESULT 7

ID HS118284 standard; RNA; EST; 430 BP.  
AC N23118;  
DT 29-DEC-1995 (Rel. 46, Created)  
DT 29-DEC-1995 (Rel. 46, Last updated, Version 1)  
DE yx67cl2.s1 Homo sapiens cDNA clone 266806 3' similar to SP:A41234  
DE A41234 MELANOCYTE-SPECIFIC PROTEIN PMEL-17 PRECURSOR - ;  
KW EST.  
OS Homo sapiens (human)  
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;  
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.  
RN [1]  
RP 1-430  
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,  
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,  
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,  
RA Trevaaskis E., Waterston R., Williamson A., Wohldmann P., Wilson R.;  
RA "The WashU-Merck EST Project";  
RL Unpublished.  
CC Contact: Wilson RK WashU-Merck EST Project Washington University  
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,  
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:  
CC est@wustl.wustl.edu High quality sequence stops: 288 Source: IMAGE  
CC Consortium, LINL This clone is available royalty-free through LINL  
CC ; contact the IMAGE Consortium (info@image.lnl.gov) for further  
CC information. NCBI gi: 1137268  
FH Key Location/Qualifiers  
FT source  
FT /organism="Homo sapiens"  
FT /clone="266806"  
FT /notes="human"  
FT mRNA  
FT <1..>430  
SQ Sequence 430 BP; 114 A; 107 G; 115 G; 93 T; 1 other;

Query Match 19.3%; Score 420; DB 130; Length 430;  
Best Local Similarity 99.3%; Pred. No. 0.00e+00;  
Matches 427; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
Db 1 caggctcgtagattatttccagtttaataagtagtctccagaggaagactggggaaatat 60  
|||||  
Cp 2132 CAGGCTCTGAGTATTATTTTCAGTTAATAGTAGTCTCCAGGGAAGACTGGGGAAATAT 2073  
|||||  
Db 61 aggtgtttctgtcaactccaggaataacacagcatcatatgagagtactcagacctgtg 120  
|||||  
Cp 2072 AGGTGTTTCTGTCAACTCCAGGAAATACAGCATCATATGACAGTACTCAGACCTGCTG 2013  
|||||  
Db 121 cccactgaggagggggtattctcaccatgggacaagagacagatcggggtagacg 180  
|||||  
Cp 2012 CCCACTGAGGAGGGGGCTGTCTCACCATGGGCAAGAGAGAGAGATGGGGGTAGACG 1953  
|||||  
Db 181 cagccagtgactgtctgtatgtgcaactgggtgacggagaagtctgtctcataagctc 240  
|||||  
Cp 1952 CAGCCAGTGACTGTCTGTATGTGGCAACTGGGTACGAGAGACTCTGTCTCATAGTCT 1893  
|||||  
Db 241 ggcctatatatcagatgcaaggacacacacacacacacacacacacacacacacacat 300  
|||||  
Cp 1892 GGCCTATATATAGAGATGCAAGGACACACGCCATCAACACCAAGATGCCACGAT 1833  
|||||  
Db 301 cagcgaaacctcccaaggctgtctctgaccagcatgataagctgggtgcttgacca 360  
|||||  
Cp 1832 CAGCGAACTGCCCAAGGCTGCTTCTTGACAGGCATGATAAGCTGGTGTCT-GACCA 1774  
|||||  
Db 361 ctgccaggtcttggatcagccagagacacattgagcgagtagtgcgccagccacct 420  
|||||  
Cp 1773 CTGCCAGCTGTGGTATGACCCAGACACATTCAGCGAGTATGTCCCGACACACCT 1714  
|||||  
Db 421 tcagttatctg 430  
|||||  
Cp 1713 TCAGTATCTG 1704  
|||||

## RESULT 8

LOCUS N23118 430 bp mRNA EST 28-DEC-1995  
DEFINITION yx67cl2.s1 Homo sapiens cDNA clone 266806 3' similar to SP:A41234  
DEFINITION A41234 MELANOCYTE-SPECIFIC PROTEIN PMEL-17 PRECURSOR - ;  
ACCESSION N23118  
NID gi137268  
KEYWORDS EST.  
SOURCE human clone=266806 primer=ml3 -40 forward library=Soares melanocyte  
2NBM vector=pT7T3D (Pharmacia) with a modified polylinker  
host=DH10B (ampicillin resistant) Reitel-Not I Reitel-Eco RI Male.  
1st strand cDNA was primed with a Not I - oligo(dT) primer  
[5'-TGTTACCAATCTGAAGTGGCGCGCGAGTTTTTTTTTTTTTTT-3'],  
double-stranded cDNA was size selected, ligated to Eco RI adapters  
(Pharmacia), digested with Not I and cloned into the Not I and Eco  
RI sites of a modified pT7T3 vector (Pharmacia). Library  
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal  
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.  
Albino.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;  
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
Sarcopterygii; Choonata; Tetrapoda; Amniota; Mammalia; Theria;  
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE  
1 (bases 1 to 430)  
AUTHORS  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,

Parsons,J., Rifkin,L., Rohlffing,T., Soares,M., Tan,F.,  
Trevaakie,E., Waterston,R., Williamson,A., Wohlmann,P. and  
Wilson,R.

TITLE  
JOURNAL  
COMMENT

The WashU-Merck EST Project  
Unpublished (1995)

Contact: Wilson RK

WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

High quality sequence stops: 288  
Source: IMAGE Consortium, LNL

This clone is available royalty-free through LNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 1137268

FEATURES  
Location/Qualifiers

1..430  
/organism="Homo sapiens"  
/clone="266806"  
/note="human"

BASE COUNT 114 a 107 c 115 g 93 t 1 others

Query Match 19.3%; Score 420; DB 113; Length 430;

Best Local Similarity 99.3%; Pred. No. 0.00e+00;

Matches 427; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Db 1 caggctctgagtatttttcagtaataagtagtctccagggaagactgggggaaatat 60

Cp 2132 CAGGCTCTGAGTATTATTTTCAGTTAATAGTAGTCTCCACGGGAAGACTGGGGGAATAT 2073

Db 61 aggtgtttctgcaactccaggaaaatacacagcatcatatgagtagtactcagacctgctg 120

Cp 2072 AGGTGTTCTCTCACTCCAGGAATACAGCATCATATGAGTACTCAGACTGCTG 2013

Db 121 cccactgaggaggggggtatttccaccaatgggacagagcagaagatcggggtagacg 180

Cp 2012 CCCACTGAGGAGGGGCTGTTCTCACTCAATGGCAAGAGCAGAGATCGGGGTAGACG 1953

Db 181 cagccagtactgctgctatgtggcaactgggtacggagaagcttctgttcataagctc 240

Cp 1952 CAGCCAGTACTGCTGCTATGTGGCACTGGGGTACGGAGAGTCTTGTCATAGTCT 1893

Db 241 ggccttatatcagagatgcaaggaccacacagccatcaacaccagcaagatgccacgat 300

Cp 1892 GGCCTATATCAGAGATGCAAGGCCACACAGCCATCAACACCAAGATGCCACGAT 1833

Db 301 cagcggaaactgccaaagcctcttctgaccaggatgataagctgggtgcttgacca 360

Cp 1832 CAGCGAACCCTGCCAAGCCTCTTCTTGACAGGCATGATAGTGGGTGCT-GACCA 1774

Db 361 ctgccagcctntgtatcagccagagacacattgaggcagtagtctcccagagcacct 420

Cp 1773 CTGCCAGCCTCTTGCTATCAGCCAGACACATTGAGGAGTATGTCGCCGAGCCACCT 1714

Db 421 tcagtatctg 430

Cp 1713 TCAGTATCTG 1704

RESULT 9

LOCUS

DEFINITION yx67al2.s1 Homo sapiens cDNA clone 266782 3' similar to SP:A41234  
A41234 MELANOCYTE-SPECIFIC PROTEIN PMEL-17 PRECURSOR - ;

ACCESSION N23111

NTID g1137261

KEYWORDS

SOURCE

human clone=266782 primer=ml3 -40 forward library=Soares melanocyte  
2NBHM vector=pf7T30 (Pharmacia) with a modified polylinker  
host=DH10B (ampicillin resistant) Reitel=Not I Reite2=Eco RI Male.

1st strand cDNA was primed with a Not I - oligo(dT) primer  
[5'-TGTTACCAATCTGAAGTGGGAGCGGCGAGTTTTTTTTTTTTTTT-3'],

double-stranded cDNA was size selected, ligated to Eco RI adapters  
(Pharmacia), digested with Not I and cloned into the Not I and Eco

RI sites of a modified pT73 vector (Pharmacia). Library  
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal

foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.  
Albino.

ORGANISM

Homo sapiens

Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;  
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;  
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 431)

Authors Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,  
Parsons,J., Rifkin,L., Rohlffing,T., Soares,M., Tan,F.,  
Trevaakie,E., Waterston,R., Williamson,A., Wohlmann,P. and  
Wilson,R.

Journal The WashU-Merck EST Project  
Unpublished (1995)

COMMENT

Contact: Wilson RK

WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

High quality sequence stops: 286

Source: IMAGE Consortium, LNL

This clone is available royalty-free through LNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 1137261

Location/Qualifiers

1..431  
/organism="Homo sapiens"  
/clone="266782"  
/note="human"

<1..>431

BASE COUNT 114 a 105 c 116 g 95 t 1 others

ORIGIN

Query Match 18.6%; Score 405; DB 113; Length 431;

Best Local Similarity 98.8%; Pred. No. 0.00e+00;

Matches 424; Conservative 0; Mismatches 2; Indels 3; Gaps 3;

Db 1 caggctctgagtatttttcagtaataagtagtctccagggaagactgggggaaatat 60

Cp 2132 CAGGCTCTGAGTATTATTTTCAGTTAATAGTAGTCTCCACGGGAAGACTGGGGGAATAT 2073

Db 61 aggtgtttctgcaactccaggaaaatacacagcatcatatgagtagtactcagacctgctg 120



Cp 2072 AGGTGTTTCTGCAACTCCAGGAAATCACAGCATCATATCAGAGTACTCAGACCTGCTG 2013  
|||||  
Db 121 cccactgaggggggtattctcccaatgggacaagagcaagatgcggggtagacg 180  
|||||  
Cp 2012 CCCACTGAGGAGGGGCTGTTCTCACCAGTGGGAGGAGGATGATCGGGGTAGACG 1953  
|||||  
Db 181 cagccagtgcctgctatgtgcaactgggtacggaggaagtcttcttcataaagtct 240  
|||||  
Cp 1952 CAGCCAGTGCCTGCTATGTGGCACTGGGTACGGAGAACTCTGTTTCATAAGTCT 1893  
|||||  
Db 241 ggcctatatatcacagatgcaaggaccacagccatcaaccagcaagatgccacgat 300  
|||||  
Cp 1892 GCGCCTATATCAGAGTCAAGGACCACAGCCATCAACAGCAGCAAGATGCCACGAT 1833  
|||||  
Db 301 cagcggaacctgcccaagccttctcttgaccagcagatgataaagtgcgtgacca 360  
|||||  
Cp 1832 CAGCGGAACCTGCCAAGGCT-GCTTCTTGACCGGCAATGAAAGTGGGTGCTGACCA 1774  
|||||  
Db 361 ctgcagcgtgtgtgatacagcagagacacatttgaggcagtgatgttccccgagccacc 420  
|||||  
Cp 1773 CT-GCCAGGCTGTGTGATCAGCCAGACACATT-GAGGCGATGTGTCGCCGAGCCACC 1716  
|||||  
Db 421 cttcagat 429  
|||||  
Cp 1715 CTTCAAGTAT 1707

RESULT 10  
ID HS111256 standard; RNA; EST; 431 BP.  
AC N23111;  
DT 29-DEC-1995 (Rel. 46, Created)  
DT 29-DEC-1995 (Rel. 46, Last updated, Version 1)  
DE yx67a12.sl Homo sapiens cDNA clone 266782 3' similar to SP:A1234  
DE A41234 MELANOCYTE-SPECIFIC PROTEIN PMEL-17 PRECURSOR -  
KW EST.  
OS Homo sapiens (human)  
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;  
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.  
RN [1]  
RP 1-431  
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,  
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,  
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,  
RA Trevasakis E., Waterston R., Williamson A., Wohldmann P., Wilson R.;  
RT "The WashU-Merck EST Project";  
RL Unpublished.  
CC Contact: Wilson RK WashU-Merck EST Project Washington University  
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,  
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:  
CC est@washington.wustl.edu High quality sequence stops: 286 Source: IMAGE  
CC Consortium, LNL This clone is available royalty-free through LNL  
CC ; contact the IMAGE Consortium (info@image.lnl.gov) for further  
CC information. NCBI gi: 1137261  
FH Key Location/Qualifiers  
EH  
FT source  
FT 1..431  
FT /organism="Homo sapiens"  
FT /clone="266782"  
FT /note="human"  
FT mRNA  
FT <1..>431  
SQ Sequence 431 BP; 114 A; 105 C; 116 G; 95 T; 1 other;  
Query Match 18.6%; Score 405; DB 129; Length 431;

Best Local Similarity 98.8%; Pred. No. 0.00e+00;  
Matches 424; Conservative 0; Mismatches 2; Indels 3; Gaps 3;  
Db 1 caggtctgagtagtatttcaatagtagtctccccaggaagaactgggggaaatat 60  
|||||  
Cp 2132 CAGGCTCTGAGTATTTATTTCACTAAATAGTAGTCTCCAGGGAAGACTGGGGAAATAT 2073  
|||||  
Db 61 aggtgtttctgtcaactccaggaaatcacagcatcatatgagtagtactcagacctgtg 120  
|||||  
Cp 2072 AGTGTCTTCTGCTCACTCCAGGAAATCACAGCATATATCAGAGTACTCAGACCTGCTG 2013  
|||||  
Db 121 cccactgaggggggtattctcaccatgggacaagagcagaagatgcggggtagacg 180  
|||||  
Cp 2012 CCCACTGAGGAGGGGCTGTTCTCAACAAATGGGCAAGCAGCAGATGCGGGGTAGACG 1953  
|||||  
Db 181 cagccagtgcctgctatgtgcaactgggtacggaggaagtcttcttcataaagtct 240  
|||||  
Cp 1952 CAGCCAGTGCCTGCTATGTGGCACTGGGTACGGAGAACTCTGTTTCATAAGTCT 1893  
|||||  
Db 241 ggcctatatatcacagatgcaaggaccacagccatcaaccagcaagatgccacgat 300  
|||||  
Cp 1892 GCGCCTATATCAGAGTCAAGGACCACAGCCATCAACAGCAGCAAGATGCCACGAT 1833  
|||||  
Db 301 cagcggaacctgcccaagccttctcttgaccagcagatgataaagtgcgtgacca 360  
|||||  
Cp 1832 CAGCGGAACCTGCCAAGGCT-GCTTCTTGACCGGCAATGAAAGTGGGTGCTGACCA 1774  
|||||  
Db 361 ctgcagcgtgtgtgatacagcagagacacatttgaggcagtgatgttccccgagccacc 420  
|||||  
Cp 1773 CT-GCCAGGCTGTGTGATCAGCCAGACACATT-GAGGCGATGTGTCGCCGAGCCACC 1716  
|||||  
Db 421 cttcagat 429  
|||||  
Cp 1715 CTTCAAGTAT 1707

RESULT 11  
LOCUS N40290 521 bp mRNA EST 22-JAN-1996  
DEFINITION yx80c07.r1 Homo sapiens cDNA clone 268044 5' similar to PIR:A53668  
A53668 glycoprotein gp100 precursor, melanocyte lineage - human ;  
ACCESSION N40290  
NID g1163835  
KEYWORDS EST.  
SOURCE human clone=268044 primer=T7 library=Soares melanocyte 2NbHM  
vector=pf7T3D (Pharmacia) with a modified polylinker host=DH10B  
(ampicillin resistant) Rsite1=Not I Rsite2=Eco RI Male. 1st strand  
cDNA was primed with a Not I - oligo(dT) primer  
[5'-TGTTACCAATCTCAAGTGGGAGCGCGAGTGTGTGTGTGTGTGTGTGTGT-3'],  
double-stranded cDNA was size selected, ligated to Eco RI adapters  
(Pharmacia), digested with Not I and cloned into the Not I and Eco  
RI sites of a modified pf7T3 vector (Pharmacia). Library  
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal  
foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.  
Albino.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;  
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;  
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 521)  
AUTHORS Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,  
Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,  
Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,  
Trevasakis E., Waterston R., Williamson A., Wohldmann P. and  
Trevasakis E., Waterston R., Williamson A., Wohldmann P. and

Wilson, R.

TITLE The WashU-Merck EST Project  
JOURNAL Unpublished (1995)  
COMMENT

Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: est@wustl.wustl.edu  
High quality sequence stops: 340  
Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.

NCBI gi: 1163835

FEATURES Location/Qualifiers

source  
1..521  
/organism="Homo sapiens"  
/clone="268044"  
/note="human"

mRNA

BASE COUNT 122 a 150 c 135 g 110 t 4 others

ORIGIN

Query Match 17.5%; Score 380; DB 118; Length 521;

Best Local Similarity 98.8%; Pred. No. 0.00e+00;

Matches 405; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

Db 1 atgacccatgtgctatctcgtggaagctgacctctctacacctgggactttggagaca 60

|||||

Qy 772 ATGACCCAGTGGCTATCTGGCTGAAGCTGACCTCTCTACACCTGGGACTTTGGAGACA 831

Db 61 gtagtgaacctgatctctcgggcaacttgggtcaactcactcacttgaagctggcc 120

|||||

Qy 832 GTAGTGGAAACCTTGATCTCTGGGCACTTTGGTGCACTACTACTTACCTGGAGCCTGGCC 891

Db 121 cagtcaactgccaggtggtcctcaggtgccattcctcactcactcctctggtcctccc 180

|||||

Qy 892 CAGTCACTGCCAGGTGCTCTCGAGGTGCGATTCTCTCAGCTCCTGTGGCTCTCTCCC 951

Db 181 cagttccaggccaccagatggcagagcccaactgcagagggccctaacacacagctg 240

|||||

Qy 952 CAGTTCCAGGCACACAGATGGGCACAGGCCAACCTGCAGAGGCCCTTAACACACAGCTG 1011

Db 241 gccagtgctactacagaagttgtgggtactacacctggtcagggcgccactgcagagc 300

|||||

Qy 1012 GCCAGTGCCTACTACAGAACTTGTGGTACTACACTGCTGGTCAAGGCCCACTCGAGACC 1071

Db 301 cctctggaaccacatctgtgcaggtgccaaacctgaagtataaagcactgcacctgtgc 360

|||||

Qy 1072 CTTCTGGNACCACTCTGTGAGGTGCCAACCACTGAGTGTATTAAGCACTGCACCTGTGC 1131

Db 361 aagaatgcccaactgcagagagcacaggttatggacacctggagaagtgcc 410

|||||

Qy 1132 AGA-TGCC-ANCTGCAGAGGACAGGATATG-ACACCTG-AGAAGGTGCC 1177

RESULT 12

ID HS290283 standard; RNA; EST; 521 BP.

AC N40290;

DT 26-JAN-1996 (Rel. 4.6, Created)

DT 26-JAN-1996 (Rel. 4.6, Last updated, Version 1)

DE yx80c07.r1 Homo sapiens cDNA clone 268044 5' similar to PIR:A53668

DE A53668 glycoprotein gp100 precursor, melanocyte lineage - human ;

KW EST.

OS Homo sapiens (human)

OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;

OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homidae.

RN [1]

RP 1-521

RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,

RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,

RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,

RA Trevaskis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;

RT "The WashU-Merck EST Project";

RL Unpublished.

CC Contact: Wilson RK WashU-Merck EST Project Washington University

CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,

CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:

CC est@wustl.wustl.edu High quality sequence stops: 340 Source: IMAGE

CC Consortium, LNL This clone is available royalty-free through LNL

CC ; contact the IMAGE Consortium (info@image.lnl.gov) for further

CC information. NCBI gi: 1163835

FH Key

FH Location/Qualifiers

FT source

1..521

/organism="Homo sapiens"

/clone="268044"

/note="human"

FT mRNA

&lt;1..&gt;521

SQ Sequence 521 BP; 122 A; 150 C; 135 G; 110 T; 4 other;

Query Match

Best Local Similarity 17.5%; Score 380; DB 131; Length 521;

Matches 405; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

Db 1 atgacccatgtgctatctcgtggaagctgacctctctacacctgggactttggagaca 60

|||||

Qy 772 ATGACCCAGTGGCTATCTGGCTGAAGCTGACCTCTCTACACCTGGGACTTTGGAGACA 831

Db 61 gtagtgaacctgatctctcgggcaacttgggtcaactcactcacttgaagctggcc 120

|||||

Qy 832 GTAGTGGAAACCTTGATCTCTGGGCACTTTGGTGCACTACTACTTACCTGGAGCCTGGCC 891

Db 121 cagtcaactgccaggtggtcctcaggtgccattcctcactcactcctctggtcctccc 180

|||||

Qy 892 CAGTCACTGCCAGGTGCTCTCGAGGTGCCATTCTCTCAGCTCCTGTGGCTCTCTCCC 951

Db 181 cagttccaggccaccagatgggcacagggcccaactgcagagggccctaacacacagctg 240

|||||

Qy 952 CAGTTCCAGGCACACAGATGGGCACAGGCCAACCTGCAGAGGCCCTTAACACACAGCTG 1011

Db 241 gccagtgctactacagaagttgtgggtactacacctggtcagggcgccaaactgcagagc 300

|||||

Qy 1012 GCCAGTGCCTACTACAGAACTTGTGGTACTACACTGCTGGTCAAGGCCCACTCGAGACC 1071

Db 301 cctctggaaccacatctgtgcaggtgccaaacctgaagtataaagcactgcacctgtgc 360

|||||

Qy 1072 CTTCTGGNACCACTCTGTGAGGTGCCAACCACTGAGTGTATTAAGCACTGCACCTGTGC 1131

Db 361 aagaatgcccaactgcagagagcacaggttatggacacctggagaagtgcc 410

|||||

Qy 1132 AGA-TGCC-ANCTGCAGAGGACAGGATATG-ACACCTG-AGAAGGTGCC 1177

RESULT 13



LOCUS R78477 416 bp mRNA EST 09-JUN-1995  
DEFINITION Y173406.r1 Homo sapiens cDNA clone 144875 3' similar to SP.A41234  
A41234 MELANOCYTE-SPECIFIC PROTEIN PMEL-17 PRECURSOR - ;  
ACCESSION R78477  
NID g854770  
KEYWORDS EST.  
SOURCE human clone=144875 library=Soares placenta Nb2HP vector-pT7T3D  
(Pharmacia) with a modified polylinker host=DH10B (ampicillin  
resistant) primer=Promega -2lml3 Rsite1=Not I Rsite2=Eco RI Female  
placenta obtained at birth (full term). 1st strand cDNA was primed  
with a Not I - oligo(dT) primer [5'  
AAGTCGAGAAATTCGGCGCCGAGGAATTTTTTTTTTTTTT 3'], double-stranded  
cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not  
I and cloned into the Not I and Eco RI sites of the modified pT7T3  
vector. Library went through one round of normalization. Library  
constructed by Bento Soares and M.Fatima Bonaldo. .  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;  
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;  
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 416)  
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,  
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,  
Trevasaki,E., Waterston,R., Williamson,A., Wohlmann,P. and  
Wilson,R.  
TITLE The WashU-Merck EST Project  
JOURNAL Unpublished (1995)  
COMMENT Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu  
High quality sequence stops: 320  
Source: IMAGE Consortium, LINL  
This clone is available royalty-free through LINL ; contact the  
IMAGE Consortium (info@image.linl.gov) for further information.  
Possible reversed clone: similarity on wrong strand  
Possible reversed clone: polyT not found.

NCBI gi: 854770

Location/Qualifiers

source 1..416  
/organism="Homo sapiens"  
/clone="144875"  
/note="human"

BASE COUNT 78 a 113 c 119 g 103 t 3 others  
ORIGIN

Query Match 16.2%; Score 351; DB 68; Length 416;  
Best Local Similarity 95.4%; Pred. No. 0.00e+00;  
Matches 397; Conservative 0; Mismatches 13; Indels 6; Gaps 6;

Db 1 agctacatccctgagctggaaggtccagatgccagctcaatctgtctacgaaagta 60  
|||||  
QY 1327 AGCTACCTATTCCTGAGCTGAGAGTCCAGATGCCAGCTCAATCATCTACGGAAGTA 1386  
|||||  
Db 61 ttacaggttcctgggcccctgctggaaggtacagccaccttaaggtggtgaagac 120  
|||||  
QY 1387 TTACAGGTTCCCTGGGCCCCCTGCTGGATGGTACAGCCACCTTAAGGCTGCTGAAGAC 1446

Db 121 aagtcocccctggattgttctgtatcgtatgtcttctccgtcacccctggacatt 180  
|||||  
QY 1447 AGTCCCCCTGGATTGTGTGTATGATATGGTTCTCTTTCCGTACCTGG-ACATT 1505  
|||||  
Db 181 gtccaggttattgaaagtcgcagatcctgcagagctgtcggttcggtgaggggatgc 240  
|||||  
QY 1506 GTCCAGGGTATT-GAAAGTCCGAGATCCTGCAGGCTGTCCGTCGGTGAGGGGATGC 1564  
|||||  
Db 241 attgagctactgtctgcacaaggcggtgcceaaaggaagcctcatgggaatct 300  
|||||  
QY 1565 ATTTCAGCTGACTGTCTCTCCAAAGGGGCTGCCAAGGAAGCCTGCATGG-AGATCT 1623  
|||||  
Db 301 catgccagggtgcagccctgccagcngctgtgccagcctgtgttaccagccca 360  
|||||  
QY 1624 CATGCCAGGTCGCAGCCCTGCCAGCGGCTGTGCAGCCTGTGCT-ACCCAGCCCA 1682  
|||||  
Db 361 gcttcgncagttggttttgcaccagatacttaagggttggttggggacattgc 416  
|||||  
QY 1683 GCCT-GCAGCTGGTTCT-GCACCAGACTGAAGGGTGGCTCGGGACATACGTC 1736  
|||||

## RESULT 14

LOCUS T29726 306 bp mRNA EST 06-SEP-1995  
DEFINITION EST92123 Homo sapiens cDNA 5' end similar to melanocyte-specific  
protein Pmel 17 (GB:M77348) (HT:1391).

ACCESSION T29726

NID g611824

KEYWORDS EST.

SOURCE human primer=M13 Reverse library=Human Skin.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;  
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;  
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 306)

## REFERENCE

## AUTHORS

Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,  
Bult,C.J., Lee,N., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,  
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W.,  
Clayton,R.A., Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,  
Fitzgerald,L.M., FitzHugh,W.M., Frichman,J.L., Geoghagen,N.S.M.,  
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle Jr,P.S.,  
Kelley,J.M., Klimek,K.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M.,  
Merrick,J.M., Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T.,  
Pellegrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L.,  
Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R.,  
Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A.,  
Coleman,T.A., Collins,E.-J., Dimke,D., Feng,P., Ferrie,A.,  
Fischer,C., Hastings,G.A., He,W.-W., Hu,J.-S., Greene,J.M.,  
Gruber,J., Hudson,P., Kim,A., Kozak,D.L., Kunsch,C., Ji,H., Li,H.,  
Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.-F., Wing,J., Xu,C.,  
Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon,M.R., Rosen,C.A.,  
Haeeltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.

Initial Assessment of Human Gene Diversity and Expression Patterns  
Based Upon 52 Million Basepairs of cDNA Sequence

## JOURNAL

Unpublished (1995)

## COMMENT

Other ESTs: THC20616

Contact: Venter, JC

The Institute for Genomic Research

932 Clopper Rd, Gaithersburg, MD 20878

Tel: 3018699056

Fax: 3018699423

Email: tdbinfo@tdb.tigr.org

For clone availability, additional sequence and expression



Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

Run on: Tue Aug 27 08:12:54 1996; MasPar time 13.38 Seconds  
436.417 Million cell updates/sec

Searched: 76839 seqs, 8833555 residues

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query		DB ID	Description	Pred. No.
		Match	Length			
1	4698	100.0	661	15	MART-1 melanoma anti	0.00e+00
2	4684	99.7	661	14	Melanoma associated a	0.00e+00
3	4679	99.6	661	15	MART-1 melanoma anti	0.00e+00
4	238	5.1	30	6	Melanoma cell surface	1.67e-08
5	206	4.4	400	1	Sequence of preprospa	3.39e-06
6	172	3.7	166	2	Intestinal mucin dedu	8.21e-04
7	176	3.7	278	2	Intestinal mucin dedu	4.35e-04
8	174	3.7	278	3	SMC-41 intestinal mu	5.98e-04
9	169	3.6	907	14	EBV gp350/220.	1.32e-03
10	169	3.6	907	3	Epstein-Barr virus (E	1.32e-03
11	165	3.5	166	3	SMC-40 intestinal mu	2.48e-03
12	162	3.4	2035	11	Host cell factor prot	3.97e-03
13	162	3.4	2035	11	Host cell factor prot	3.97e-03

RESULT	1
ID	R84854 standard; Protein; 661 AA.
AC	R84854;
DT	08-MAY-1996 (revised)
DT	20-APR-1996 (first entry)
DE	MART-1 melanoma antigen cDNA25.
KW	cDNA25; MART-1; melanoma antigen recognised by T-cell;
KW	gp100 antigen derivative; melanoma; metastatic melanoma;
KW	tumour-associated antigen; immunogen; diagnosis; prognosis;
KW	prophylaxis; therapy; vaccine.
OS	Mammalian sp.

## Location/Qualifiers

ET Peptide 457..466  
 FT /label= antigenic\_peptide  
 FT /note= "see R84199"  
 PN W09529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; W05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-411714.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR MPI; 95-382963/49.  
 DR N-PSDB; 702716.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.

PS Claim 81; Fig 5A; 184pp; English.  
 CC cDNA2 is a melanoma antigen (MART-1) which is recognized by  
 CC T-lymphocytes, and is a derivative of the melanocyte-melanoma-  
 CC specific antigen gp100 (see R84855). Antigen cDNA25 is a source  
 CC of immunogenic peptides (see R84199) which are optionally modified  
 CC (see R84200-R84211) to enhance their binding to a MHC molecule and  
 CC used in medicaments, especially vaccines, for the treatment or  
 CC prevention (by immunisation) of melanoma. Antibodies against cDNA2  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of the antigen from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 661 AA;

Query Match 100.0%; Score 4698; DB 15; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
 Matches 661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mdlvklrcllhavlgallavatkvpnqdlgvsrglrltkawnrqlypewteagrlcdc 60  
 QY 1 MDLVKRLCLLHVLAVIGALLAVGATKVPNNQDLGVSRLRTKAWNRQLYPEWTEAQRDC 60  
 Db 61 wrqgqvsikvnsdptliganasfialnfpqskvlpdgqviwvntlingsqvvggqp 120  
 QY 61 WRGQVSILKVSNDGPTLIGANASFSIALNFPQSKVLPDQGVIVWNTLINGSQVWGQP 120  
 Db 121 vvpqetddaciipdgpcpgagswskrsfvyvkwgqvgflgpgvsglsgtgramlg 180  
 QY 121 VYPQETDDACIIPDGGPCPGAGSWSKRSFYVYVWKTWGMQYMQFLGGPVSGLSIGTGRAMLG 180  
 Db 181 thtmvtyvhrgrsvyvpplahessafitdqvpsvevseqlraldgnkhflrncpltf 240  
 QY 181 THTMVTVVHRGRSGRSTVPPLAHSSAFITDQVPFSVSQRLALDGNKHFLRNQPLTF 240  
 Db 241 alqlhdpqsgylaeadlsytwdfcdsgstllsralvvtthylepgvtaqvllaaplt 300  
 QY 241 ALQLHDPSCYLAEDLSYTWDFCDSSGTLISRALVVTHYLEPGVTAQVVLQAAPLTS 300  
 Db 301 cgspsvpgttdghrptaeapnttagvpttevgttppqaptaesgtsvqvpptevis 360  
 QY 301 CGSSPVPCTTDGHRPTAEAPNTTAGQVPTTEVGTTPQAPTAESPSCGTSVQVPTTEVIS 360  
 Db 361 tapvqmpataestgmtpekvpvsevmgttlaemstpeatgmtpaevsiivlsgttaqvtt 420  
 QY 361 TAPVQMPATAESTGMTPEKVPVSEVMGTTLAEMSTPEATGMTPAEVSIVVLSGTTAAQVTT 420  
 Db 421 tewettarelpipepgcdassimstesitgslgplldgtatrlrvkrqvpidevlyry 480  
 QY 421 TEWETTARELPPEPGCDASSIMSTESITGSLGPLLDGTATRLRVKRQVPIDCVLYRY 480  
 Db 481 gsfvrtldivqgieaealqavpsgedafeltvcqgglpkacameispgcqpqr1 540  
 QY 481 GSFVRTLDIVQIEAEALQAVPSGEDAFELTVSCQGLPKACAMEISPGCQPQRL 540  
 Db 541 cqpvlpepacqlvhlqkgegtcynlvaladtnslavstqlimpqgeaglgvpliv 600  
 QY 541 CQPVLPSACQLVHLQKGSCTCYNLVALADTNSLAVSTQLIMPQGEAGLGVPILV 600  
 Db 601 gillvmaavlasliyrllmkqdfsvpqlphesshlrlprifcscplgenspllsqgq 660  
 QY 601 GILLVMAVVASLILYRRLMKQDFSVPLPHSSHWLRLPRIFCSCPLGENSPILSQGQ 660  
 Db 661 V 661  
 QY 661 V 661

RESULT 2  
 ID R78646 standard; Protein; 661 AA.  
 AC R78646;  
 DT 22-JAN-1996 (first entry)  
 DE Melanoma associated antigen gp100.  
 KW Melanoma; antigen; vaccine; immunogen; primer; probe; detection;  
 KW identification; tumour; gp100.  
 OS Homo sapiens.  
 PN EP-668350-A1.  
 PD 23-AUG-1995.  
 PF 14-FEB-1995; 200348.  
 PR 16-FEB-1994; EP-200337.  
 PR 21-DEC-1994; EP-203709.  
 PA (AIKU ) AKZO NOBEL NV.  
 PI Adema GU, Figdor CG;  
 DR WPI; 95-284790/38.  
 PT Melanoma associated antigen gp100 - used in vaccines and for the  
 PS Claim 1; Page 22-24; 40pp; English.  
 CC Immunogenic peptides derived from the melanoma associated antigen  
 CC may be used in the production of vaccines. Nucleotide sequences  
 CC encoding the immunogenic peptides may be used as primers and probes  
 CC in the detection of melanoma cells. Tumour infiltrating lymphocytes  
 CC capable of binding to the melanoma associated antigen can be  
 CC cultured ex vivo and returned to melanoma particles, and when  
 CC radiolabelled, they may be used to identify tumour deposits.  
 SQ Sequence 661 AA;

Query Match 99.7%; Score 4684; DB 14; Length 661;  
 Best Local Similarity 99.8%; Pred. No. 0.00e+00;  
 Matches 660; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 mdlvklrcllhavlgallavatkvpnqdlgvsrglrltkawnrqlypewteagrlcdc 60  
 QY 1 MDLVKRLCLLHVLAVIGALLAVGATKVPNNQDLGVSRLRTKAWNRQLYPEWTEAQRDC 60  
 Db 61 wrqgqvsikvnsdptliganasfialnfpqskvlpdgqviwvntlingsqvvggqp 120  
 QY 61 WRGQVSILKVSNDGPTLIGANASFSIALNFPQSKVLPDQGVIVWNTLINGSQVWGQP 120  
 Db 121 vvpqetddaciipdgpcpgagswskrsfvyvkwgqvgflgpgvsglsgtgramlg 180  
 QY 121 VYPQETDDACIIPDGGPCPGAGSWSKRSFYVYVWKTWGMQYMQFLGGPVSGLSIGTGRAMLG 180  
 Db 181 thtmvtyvhrgrsvyvpplahessafitdqvpsvevseqlraldgnkhflrncpltf 240  
 QY 181 THTMVTVVHRGRSGRSTVPPLAHSSAFITDQVPFSVSQRLALDGNKHFLRNQPLTF 240  
 Db 241 alqlhdpqsgylaeadlsytwdfcdsgstllsralvvtthylepgvtaqvllaaplt 300  
 QY 241 ALQLHDPSCYLAEDLSYTWDFCDSSGTLISRALVVTHYLEPGVTAQVVLQAAPLTS 300  
 Db 301 cgspsvpgttdghrptaeapnttagvpttevgttppqaptaesgtsvqvpptevis 360  
 QY 301 CGSSPVPCTTDGHRPTAEAPNTTAGQVPTTEVGTTPQAPTAESPSCGTSVQVPTTEVIS 360  
 Db 361 tapvqmpataestgmtpekvpvsevmgttlaemstpeatgmtpaevsiivlsgttaqvtt 420  
 QY 361 TAPVQMPATAESTGMTPEKVPVSEVMGTTLAEMSTPEATGMTPAEVSIVVLSGTTAAQVTT 420  
 Db 421 tewettarelpipepgcdassimstesitgslgplldgtatrlrvkrqvpidevlyry 480

Qy 421 TEWVETARELPPEPEGDASSIMSTESITGSLGLDGTATLRLVKRQVPLDCVLYR 480  
 Db 481 gsfvltldivqgisaeilqavpsgegdafeltvscggglpkeacmeispgcqpqr1 540  
 Qy 481 GSFVTLTDIVQGISAEILQAVPSGEGDAFELTVSCGGGLPKEACMETSSPGCQPPAQR1 540  
 Db 541 cqvlpsepacqlvhlqilkgsgtycinvladtnslavvstqlimpqgeaglgqvpliv 600  
 Qy 541 CQVLPSPACQLVHLQILKGSGTYCLNVLADTNSLAVVSTQLIMPQGEAGLGQVPLIV 600  
 Db 601 gillivmavvlasliyrllmkqdfsvqqlphssshwlrprifescpiqenapllsqgq 660  
 Qy 601 GILLVIMAVVLASLIYRRLMKQDFSVQQLPHSSSHWLRLPRIFCSCPIGENSPILLSGQQ 660  
 Db 661 v 661  
 Qy 661 V 661

RESULT 3  
 -ID R84855 standard; Protein; 661 AA.  
 AC R84855;  
 DT 08-MAY-1996 (revised)  
 DT 20-APR-1996 (first entry)  
 DE MART-1 melanoma antigen gp100.  
 KW gp100; MART-1; melanoma antigen recognised by T-cell;  
 KW cDNA25 antigen derivative; melanocyte; melanoma;  
 KW metastatic melanoma; tumour-associated antigen; immunogen;  
 KW diagnosis; prognosis; prophylaxis; therapy; vaccine.  
 OS Mammalian sp.

FH Key Location/Qualifiers  
 FT Peptide 154..163  
 FT /label= C9-154 immunogenic\_peptide  
 FT Peptide 208..217  
 FT /label= C9-209 immunogenic\_peptide  
 FT /note= "see R84210"  
 FT Peptide 280..288  
 FT /label= C9-280 immunogenic\_peptide  
 FT /note= "see R84208"  
 FT Peptide 457..266  
 FT /label= immunogenic\_peptide  
 FT Peptide 476..485  
 FT /label= immunogenic\_peptide  
 PN W09529193-A2.  
 PD 02-NOV-1995.  
 PE 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectore, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 81; Fig 7A; 184pp; English.  
 CC gp100 is a melanocyte-melanoma-specific antigen (MART-1) which  
 CC is recognized by T-lymphocytes, and is a derivative of the  
 CC melanoma-specific antigen cDNA25 (see R84854). gp100 is a source  
 CC of immunogenic peptides which are optionally modified to enhance  
 CC their binding to a MHC molecule, and used in medicaments,  
 CC especially vaccines, for the treatment or prevention (by  
 CC immunisation) of melanoma. Antibodies against cDNA2 and its  
 CC immunogenic peptides may be used in the detection and isolation

CC of the antigen from a sample, the detection of which is indicative  
 CC of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 661 AA;

Query Match 99.6%; Score 4679; DB 15; Length 661;  
 Best Local Similarity 99.7%; Pred. No. 0.00e+00;  
 Matches 659; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 mdlvlkrcllhlavigallavgatkvprnqdlwvsrqlrtkawnrqlypewteaqrldc 60  
 Qy 1 MDLVLRCLLHLAVIGALLAVGATKVPRNQDMLGVSRLQRTKAWNRQLYPEWTEAQR1DC 60  
 Db 61 wrggqvskvnsdgtlignasfsialnfpsqskvlpdgqvlwvnnntiingeqvqgqp 120  
 Qy 61 WRGGQVSKVNSDGTLLIGNASFIALNFPQSKVLPDQGVLIWVNNNTIINGSQVWGGQP 120  
 Db 121 vvpqetddacifpdgpcpsqswskrfvyvkwkwgqvwlpgpvgslgiqvgamlg 180  
 Qy 121 VYPQETDDACIFPDGPCPSQSWSKRFVYVWKWKGQVWLPGPVGSLGIQVGRAMLG 180  
 Db 181 thtmvtyvhrgrsryvplahssafitdqvpfsveqqlraldgnkhflrnqpltf 240  
 Qy 181 THTMVTVYHRGRSRYVPLAHSSSAFTIQVPFSVSVQLRALDGNKHF1LRNQPLTF 240  
 Db 241 alqlhdpssylaeadlsytwdfgsotllsralvvthtylepqgqvtaqvqlqaapls 300  
 Qy 241 ALQLHDPSSYLAEADLSYTWDFGSOTLLSRA1LVVTHTYLEPGQVTAQVVLQAALPLTS 300  
 Db 301 cgsfpvgttddhrtaeapntaagvptteevgttgcqtaepsqtsqvpttevis 360  
 Qy 301 CGSFPVGTDDGHRPTAEAPNTAGQVPTTEVVGTTCQAPTAEPSGTTSVQVPTTEVIS 360  
 Db 361 tapvqmpataestgmtpekvpsvevmgtllaemstpeatgmtpaevsiivlsqtaqvtt 420  
 Qy 361 TAPVQMPATAESTGMTPEKVPSEVMGTTLAEMSTPEATGMTPAEVS1VWLSGTTAAQVTT 420  
 Db 421 tewvettarelpipepeqpdaesimstesitgslgplldgtatrlrvkrqpldcvlyr 480  
 Qy 421 TEWVETARELPPEPEGDASSIMSTESITGSLGLDGTATLRLVKRQVPLDCVLYR 480  
 Db 481 gsfvltldivqgisaeilqavpsgegdafeltvscggglpkeacmeispgcqpqr1 540  
 Qy 481 GSFVTLTDIVQGISAEILQAVPSGEGDAFELTVSCGGGLPKEACMETSSPGCQPPAQR1 540  
 Db 541 cqvlpsepacqlvhlqilkgsgtycinvladtnslavvstqlimpqgeaglgqvpliv 600  
 Qy 541 CQVLPSPACQLVHLQILKGSGTYCLNVLADTNSLAVVSTQLIMPQGEAGLGQVPLIV 600  
 Db 601 gillivmavvlasliyrllmkqdfsvqqlphssshwlrprifescpiqenapllsqgq 660  
 Qy 601 GILLVIMAVVLASLIYRRLMKQDFSVQQLPHSSSHWLRLPRIFCSCPIGENSPILLSGQQ 660  
 Db 661 v 661  
 Qy 661 V 661

RESULT 4  
 ID R30054 standard; peptide; 30 AA.  
 AC R30054;  
 DT 05-MAY-1993 (first entry)  
 DE Melanoma cell surface protein N-terminal sequence.  
 KW N-terminal; membrane; bound; ME20; antigen; human; melanoma; tumour;  
 KW antibody; immune response.







```
Query Match      3.7%; Score 174; DB 3; Length 278;
Best Local Similarity 28.0%; Pred. No. 5.98e-04;
Matches 46; Conservative 40; Mismatches 70; Indels 8; Gaps 6;

Db      6 aptpistttttvptptgtgtptttttttttttvptptgt---qtpptvlltttt 62
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     304 SPVCTTGDGHRPTAENNTAGQVPTTEVGTTPGQAPTAEPSTGTSVQVPTTEVISTAP 363
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db      63 tmtptptstktvt-pittttttvtrtpgt-qtpmipistttttvptptpigs 120
      ||: ||: : : : : ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||
QY     364 VQMPSTAESTGMPKVPVSEVMGTTLAEMSTPEATGMPAEVSIIVLSGTTAAQVTTTEW 423
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db      121 tgpptstapiaelttntppesetpqtarstsepltesttll 164
      : : ||: ||: : : : : ||: ||: ||: ||: ||: ||: ||: ||: ||
QY     424 VET-T-ARELPPEPEGDASSIMST-ESITGSLGPLLDGTA-TL 464

RESULT 9
ID R80144 standard; Protein; 907 AA.
AC R80144;
DT 18-JAN-1996 (first entry)
DE EBV gp350/220.
KW EBV; gp350; gp220; gp350/gp220; non-splicing variant; vaccine.
OS Epstein-Barr virus.
FH Key Location/Qualifiers
FT Peptide 1..18
FT /label= Sig_peptide
FT Region 861..881
FT /note= "transmembrane region"
PN W09528488-A1.
PD 26-OCT-1995.
PF 13-APR-1995; U04611.
PR 18-APR-1994; US-229791.
PA (AVIR-) AVIRON.
PI Jackman WT, Spaete R;
DR N-PSDB; T04821.
PT New DNA encoding a homogeneous gp350 protein - can be used for
PT preventing and treating Epstein-Barr virus-related diseases or
PT conditions
PS Disclosure; Fig.1; 6lpp; English.
CC The donor and acceptor splice sites of the EBV gene encoding gp350/
CC 220 are mutated by replacement of native nucleotides by non-native
CC nucleotides, without altering the encoded amino acid sequence,
CC resulting in elimination of gp220 prodn. Recombinant homogeneous
CC gp350, useful in vaccines, is expressed in mammalian or insect cell
CC hosts.
SQ Sequence 907 AA;

Query Match      3.6%; Score 169; DB 14; Length 907;
Best Local Similarity 23.2%; Pred. No. 1.32e-03;
Matches 49; Conservative 56; Mismatches 97; Indels 9; Gaps 9;

Db      421 fkapeatstptlntgfdapntttgtpsthvptnltpaetgptvstadvtsptpag 480
      |: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     262 FGDSSGTLISRALVTHYLEPQVTA-QVVLQAALPLTSCGSS-PVPCTDGHRTAE 319
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db      481 ttsgasvtpspwdngtsgkapdm-tsstspvttptnatpsptpavttptnatpsptp 539
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     320 PNNTAGQV-PTTEV-VGTTGQAPTAEPSTGTSVQVPTTEVISTAP-VQMPSTAESTG 376
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db      540 avttptnatpsptlgktspsavttptnatpsptlgktspsavttptnatpsptlgkts 599
      |: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     377 E-KVPVSEVMGTTLAEMSTPEATGMPAEVSIIVLSGTTAAQVTTTEWETTARELP 435
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Query Match      3.6%; Score 166 AA.
Best Local Similarity 23.2%; Pred. No. 1.32e-03;
Matches 49; Conservative 56; Mismatches 97; Indels 9; Gaps 9;

Db      421 fkapeatstptlntgfdapntttgtpsthvptnltpaetgptvstadvtsptpag 480
      |: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     262 FGDSSGTLISRALVTHYLEPQVTA-QVVLQAALPLTSCGSS-PVPCTDGHRTAE 319
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db      481 ttsgasvtpspwdngtsgkapdm-tsstspvttptnatpsptpavttptnatpsptp 539
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     320 PNNTAGQV-PTTEV-VGTTGQAPTAEPSTGTSVQVPTTEVISTAP-VQMPSTAESTG 376
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db      540 avttptnatpsptlgktspsavttptnatpsptlgktspsavttptnatpsptlgkts 599
      |: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     377 E-KVPVSEVMGTTLAEMSTPEATGMPAEVSIIVLSGTTAAQVTTTEWETTARELP 435
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Query Match      3.7%; Score 174; DB 3; Length 278;
Best Local Similarity 28.0%; Pred. No. 5.98e-04;
Matches 46; Conservative 40; Mismatches 70; Indels 8; Gaps 6;

Db      6 aptpistttttvptptgtgtptttttttttttvptptgt---qtpptvlltttt 62
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     304 SPVCTTGDGHRPTAENNTAGQVPTTEVGTTPGQAPTAEPSTGTSVQVPTTEVISTAP 363
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db      63 tmtptptstktvt-pittttttvtrtpgt-qtpmipistttttvptptpigs 120
      ||: ||: : : : : ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||
QY     364 VQMPSTAESTGMPKVPVSEVMGTTLAEMSTPEATGMPAEVSIIVLSGTTAAQVTTTEW 423
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db      121 tgpptstapiaelttntppesetpqtarstsepltesttll 164
      : : ||: ||: : : : : ||: ||: ||: ||: ||: ||: ||: ||: ||
QY     424 VET-T-ARELPPEPEGDASSIMST-ESITGSLGPLLDGTA-TL 464
```

```
Db      600 ptaavttptnatgptvgtspqanathl 630
      |: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     436 PEGPDASSI-MSTESITGSLGPLLDGTA-TL 464

RESULT 10
ID P50073 standard; Protein; 907 AA.
AC P50073;
DT 17-OCT-1991 (first entry)
DE Epstein-Barr virus (EBV) outer surface protein.
KW Epstein-Barr virus; antigen; vaccine.
OS Epstein-Barr virus.
PN EP-151079-A.
PD 07-AUG-1985.
PF 28-JAN-1985; 400141.
PR 30-JAN-1984; US-575352.
PR 23-JUL-1984; US-633558.
PA (UYCH-) UNIV OF CHICAGO.
PI Kleff E, Tanner J, Hummel M, Belseel C.
DR MPI; 85-191978/32.
DR N-PSDB; N50114.
PT New fragment of Epstein-Barr Virus DNA - useful in vector to
PT express polypeptide for use in prepn. of vaccine against the
PT virus and for use in diagnosis.
PS Claim 2; Page 23-25; 26pp; English.
CC The sequence encodes an outer surface viral protein of EBV, used
CC to generate antibodies reacting with the surface proteins of
CC EBV-infected cells, and in the preparation of a vaccine against EBV.
SQ Sequence 907 AA;

Query Match      3.6%; Score 169; DB 3; Length 907;
Best Local Similarity 23.2%; Pred. No. 1.32e-03;
Matches 49; Conservative 56; Mismatches 97; Indels 9; Gaps 9;

Db      421 fkapeatstptlntgfdapntttgtpsthvptnltpaetgptvstadvtsptpag 480
      |: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     262 FGDSSGTLISRALVTHYLEPQVTA-QVVLQAALPLTSCGSS-PVPCTDGHRTAE 319
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db      481 ttsgasvtpspwdngtsgkapdm-tsstspvttptnatpsptpavttptnatpsptp 539
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     320 PNNTAGQV-PTTEV-VGTTGQAPTAEPSTGTSVQVPTTEVISTAP-VQMPSTAESTG 376
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db      540 avttptnatpsptlgktspsavttptnatpsptlgktspsavttptnatpsptlgkts 599
      |: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     377 E-KVPVSEVMGTTLAEMSTPEATGMPAEVSIIVLSGTTAAQVTTTEWETTARELP 435
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db      600 ptaavttptnatgptvgtspqanathl 630
      |: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     436 PEGPDASSI-MSTESITGSLGPLLDGTA-TL 464

RESULT 11
ID R12534 standard; Protein; 166 AA.
AC R12534;
DT 02-SEP-1991 (first entry)
DE SMUC-40 intestinal mucin.
KW SMUC; intestinal mucin; diagnosis; cancer.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Region 3..25
FT /label= repeat_1
FT Region 26..48
FT /label= repeat_2
```

Query Match	3.5%;	Score 165;	DB 3;	Length 166;
Best Local Similarity	28.5%;	Pred. No.	2.48e-03;	
Matches	45;	Conservative	35;	Mismatches 69; Indels 9; Gaps 9;

Db	9	stttmvtpttptgtgctllprpirtttttvtp-tg-tqpttatsttttvtp	tptp 66
	:	l:::	:     :   :   :   :   :
QY	265	SSGTLISRALVWHT-YLEPGV-TAQWLQAAILPLSCGSFPTDIGHRP	TAEAPNT 322
	:	:	:
Db	67	tgtqtptttittttttvtpptgtgt-ptptp-itttttv-tptptpg-tqtp	tptp 122
	:	:       :     :     :     :     :     :     :	
QY	323	TAGQVPTTEWVGTPGOAPTAEPSGTTSVQVPTTETSTAPVQMPTAESTGM	TEKVPVS 382
	:	:	:
Db	123	itttttvtpptpgt-qtpstatpitnttvtp	tptp 159
	:	::    :     :     :     :	
QY	383	EVMTGTTIAEMSTPEATGMTPAECSIVVLSCGTTAAQVTT	420

PF 03-DEC-1993; U11721.  
PR 04-DEC-1992; US-989842.  
PR 12-APR-1993; US-046585.  
PA (COLD-) COLD SPRING HARBOR LAB.  
PA (TULA-) TULARIK INC.  
PI Herr W, Lamaco K, Wilson A;  
PI WPI; 94-234207/28.  
DR N-PSDB; Q69229.  
DR  
PT New Host Cell Factor polypeptide(s) and nucleic acid - are used  
PT to develop agents for diagnosis or treatment of disease  
PT associated with expression of a HCF-modulated gene e.g. viral

FT	/note= "peptide 860"	
FT	Active site	168..186
FT	/note= "peptide 837"	
FT	Active site	333..340
FT	/note= "peptide 852"	
FT	Active site	426..449
FT	/note= "peptide 362"	
FT	Active site	511..526
FT	/note= "peptide 329"	
FT	Active site	578..594
FT	/note= "peptide 223 first sequence"	
FT	Active site	594..611
FT	/note= "peptide 826 1st peptide"	
FT	Active site	611..623
FT	/note= "peptide 223 2nd sequence"	
FT	Active site	723..731
FT	/note= "peptide 318"	
FT	Active site	802..813
FT	/note= "peptide 299"	
FT	Active site	813..820
FT	/note= "peptide 268"	
FT	Active site	836..847
FT	/note= "peptide 826 2nd sequence"	
FT	Active site	1010..1031
FT	/note= "THE TNT repeat 1"	
FT	Active site	1072..1093
FT	/note= "THE TNT repeat 2"	
FT	Active site	1101..1126
FT	/note= "THE TNT repeat 3"	
FT	Active site	1158..1183
FT	/note= "THE TNT repeat 4"	
FT	Active site	1286..1311
FT	/note= "THE TNT repeat 5"	
FT	Active site	1314..1339
FT	/note= "THE TNT repeat 6"	
FT	Active site	1349..1374
FT	/note= "THE TNT repeat 7"	
FT	Active site	1414..1439
FT	/note= "THE TNT repeat 8"	
FT	Active site	1774..1781
FT	/note= "peptide 293 2nd sequence"	
FT	Active site	1808..1819
FT	/note= "peptide 115"	
FT	Active site	1819..1840
FT	/note= "peptide 261 1st sequence"	
FT	Active site	1853..1863
FT	/note= "peptide 240"	
FT	Active site	1901..1919
FT	/note= "peptide 832"	
FT	Active site	1919..1930
FT	/note= "peptide 261 2nd sequence"	
PN	W09413315-A.	
PD	23-JUN-1994.	
PR	03-DEC-1993; U11721.	
PR	04-DEC-1992; US-989842.	
PR	12-APR-1993; US-046585.	
PA	(COLD)- COLD SPRING HARBOR LAB.	
PA	(TULA)- TULARIK INC.	
PI	Herr W, Lamarco K, Wilson A;	
DR	WPI; 94-234207/28.	
DR	N-PSDB; Q69229.	
PT	New Host Cell Factor polypeptide(s)	
PT	to develop agents for diagnosis or	
PT	associated with expression of a HCF	

Query Match	3.3%;	Score 157;	DB 10;	Length 252;
Best Local Similarity	26.0%;	Pred. No. 8.6e-03;		
Matches	39;	Conservative 43;	Mismatches 58;	Indels 7; Gaps 7;
Ddb	46	aattvattvvvtmaa-aattvvaatmavaaatvva-aattvav-ttmavaattv	102	
	:	:	:	:
	:	:	:	:
	::	::	::	::
	:	:	:	:
	:	:	:	:
Oy	285	PVTAQVVQLAAIPLTSCGSSPVPCTTDGHRPTAEAPNTTAGOVPTTEVGVTTCQAPTAEE	344	
Ddb	103	vaatnavav-tttvvaatmavaaatvvaatmavaalittvvaatmaatmvimaettatt-maat	160	

Query Match	3.2%;	Score 148;	DB 15;	Length 671;
Best Local Similarity	28.2%;	Pred. No. 3.48e-02;		
Matches	55;	Conservative 49;	Mismatches 76;	Indels 15;
Gaps	13;			
Db	27	sttlfsealivsvgtpvallpvtaeeteqptn--aevaappttetglvetpttettggi 85		
		: :        : : :   :   :        :		
Qy	266	SGTLISRALVVTTH--TYLEPGVTAQWVQAIAPLTSCGSSVPVCTTDGHRPTAEAPNTT 323		
Db	86	teq-ptidsstttesttessketpttpsteqptvd--sttpvesgttdss--va-eitpva 140		
		:   :   :   :   :   :   :   :   :   :   :   :   :   :		
Qy	324	AQGVPTTEWGTTPGQAPTA-EPSGTTSVQVPTTEVISTAPVQMPAESTGMTKEKVPVS 382		
Db	141	pttteseaaaptddkvvpearvasadtfsalsptq--spaeifaelarcaaqpiaqand 199		
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :		
Qy	383	-EVMGTTTIAEMSTPEATGMTP-AEVSIV-VLSGTTAAQVTTTWEVETAR-ELIPIPEG 438		
Db	200	lyas--vmaagives 213		
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :		
Qy	439	PDASSIMSTESITGS 453		

RESULT	15
ID	R75107 standard; Protein; 722 AA.
AC	R75107;

Aug 27 08:01

US-08-231-565A-27.rag

17

DT 01-FEB-1996 (first entry)  
 DE Bile salt-activated lipase with 16 C-terminal 11-mer repeats.  
 KW Bile salt activated lipase; BAL; C-terminal repeat.  
 OS Mammalia.  
 FH Key Location/Qualifiers  
 FT Region 539..722  
 FT /note= "contains 16 copies of imperfect 11mer  
 FT repeats; the region is made up from a  
 FT combination of 7 different sequences"  
 PN J07111891-A.  
 PD 02-MAY-1995.  
 PF 30-SEP-1993; 245079.  
 PR 30-SEP-1993; JP-245079.  
 PA (MEIP ) MEIJI MILK PROD CO LTD.  
 DR WPI; 95-196318/26.  
 DR N-PSDB; Q90579.  
 PT New recombinant bile salt-activated lipase(s), DNA encoding them  
 PT and vectors - useful for high level expression of the lipase(s) by  
 PT fermenter-culturing.  
 PS Disclosure; Page 10-13; 31pp; Japanese.  
 CC R75099-R75105 (encoded by Q90570-Q90576) are attached to the C-  
 CC terminus of the bile salt-activated lipase (BAL) of sequence R75098  
 CC (encoded by Q90569) and/or in which the amino acid sequence SMTGS  
 CC (encoded by Q90578) is attached to the N-terminus of the BAL. The  
 CC recombinant lipases can be produced by fermentation in Pichia  
 CC pastoris GS 115 transformed host cells. The present sequence  
 CC is a protein with 16 copies of the undecapeptide imperfect  
 CC repeats at the BAL C-terminus.  
 SQ Sequence 722 AA;

Query Match 3.1%; Score 147; DB 14; Length 722;  
 Best Local Similarity 24.6%; Pred. No. 4.05e-02;  
 Matches 42; Conservative 50; Mismatches 67; Indels 12; Gaps 12;  
 Db 547 eatpvptgdsapvpptgdsagppvptgdsagppvptgdsagppv 606  
 Qy 282 EFGPVTAQVVLQAA-IPLTS-CGSSPVPGTTD-GHRPTAENPTTAGQVPTTEVGTTPG 338  
 Db 607 -pvtgd-sgappvp-ptgds-gappvp-ptgds-g-appvpptgdagppvpptgdsagp 659  
 Qy 339 QAPTAEFGSTTSVQVPTTEVISTAPVQMPTEASTGMPTEKVPVSEVMGTTLAEMSTPEAT 398  
 Db 660 pvpptgdsagppvptgdsapvpptgdsag-appvpptgdsagappvpptd 709  
 Qy 399 GMTPA-EVSIVVLSGTTAAQVTTTEWVETTTARELP IPEECPDASSIMSTE 448

Search completed: Tue Aug 27 08:14:12 1996  
 Job time : 78 secs.

Result	No.	Score	Query		ID	Description	Pred. No.
			Match	Length			
1	4684	99.7	661.11	A53668	glycoprotein gp100 p	0.00e+00	
2	4595	97.8	626.12	A1234	melanocyte-specific	0.00e+00	
3	3004	63.9	626.12	S53871	Pmel 17 protein - mo	0.00e+00	
4	2541	54.1	491.12	A49179	melanoma antigen hom	0.00e+00	
5	217	4.6	873.14	A47283	calphotin - Drosophi	8.01e-11	
6	212	4.5	294.11	A37232	mucin, tracheal (AMN	2.97e-10	
7	206	4.4	400	5 A28172	spasmolysin precurs	1.41e-09	
8	205	4.4	865.10	A47282	calcium-binding prot	1.83e-09	
9	200	4.3	447.12	A39321	mucin - rat (fragmen	6.63e-09	
10	198	4.2	322.12	A53715	apomucin precursor -	1.11e-08	
11	183	3.9	279.11	S53363	mucin JER58 - huma	4.91e-07	
12	175	3.7	167.11	A33532	mucin VM38-40 - huma	3.56e-06	

## ALIGNMENTS

RESULT	1
ENTRY	A53668
TITLE	#type complete
ALTERNATE_NAMES	glycoprotein gp100 precursor, melanocyte lineage - human melanoma antigen 25
ORGANISM	#formal name Homo sapiens #common name man
DATE	07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 26-May-1995
ACCESSIONS	A53668; A55753
REFERENCE	A53668
#authors	Adema, G.J.; de Boer, A.J.; Vogel, A.M.; Loenen, W.A.M.; Fidor, C.G.
#journal	J. Biol. Chem. (1994) 269:20126-20133
#title	Molecular characterization of the melanocyte lineage-specific antigen gp100.

```
##status      not compared with conceptual translation
```

```
##molecule type mRNA
##residues 1-161, F', 163-661 ##label KAW
##note nucleotide sequence not given
KEYWORDS glycoprotein
SUMMARY #length 661 #molecular-weight 70255 #checksum 5487

Query Match 99.7%; Score 4684; DB 11; Length 661;
Best Local Similarity 99.8%; Pred. No. 0.00e+00;
Matches 660; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 mdlvklrclhlhavigallavgatkvprnqdlgvsrqlrtkawnrqlypewteaqrldc 60
Qy 1 MDLVKRCLLHLAVIGALLAVGATKVP RNQDLGVS RQLRTKAWNRQLYP EWTEAQR LDC 60

Db 61 wrggqvskvndgptlliganasfsialnfpqskvlpdgqvlwvntlingsqvwgqp 120
Qy 61 WRGGQVSKVNDGPTLLIGANASFSIALNFPQSKVLPDGQVLWVNTLINGSQVWGQP 120

Db 121 vypeqtdacifpdgpcpsqgswqkrfvyvkwkgvqvlpgpvgaglsigtgramlg 180
Qy 121 VYPEQTDACIFPDGPCPSQGSWQKRFVYVWKWKGVLPGPVGAGLSIGTGRAMLG 180

Db 181 thtmvtyvhrgrsyrpplahsaeafitcdqvpfsveqlraldgnkhflrnpqltf 240
Qy 181 THTMVTVVHRGRSRYPP LAHSSAEFTITDQVPFSVSQLRALDGNKHFLRNPQLTF 240

Db 241 alqlhdpsglaeaadleytwdfgdsqgtllsralvvtthylepqpvtqavqlaaipits 300
Qy 241 ALQLHDP SGYLA EADLSYTWDFGDSGTLLSRA LVVTHYLEPQPVTQAVQLAAIPTS 300

Db 301 cgssvpvgttddghrptaeapnttaqvpvtevgtttqgaptaeapsgttsqvpttevis 360
Qy 301 CGSSVPVGTDDGHRPTAEAPNTTAQVPVTEVGTTPQAPTAEPSGTTSVQVPTTEVIS 360

Db 361 tapvqmpataestgmtpekvpvsevmgtllaemstpeatgmtpaevsiwvlgtaaqrvt 420
Qy 361 TAPVQMPATAESTGMTPEKVPVSEVMGTTLAEMSTPEATGMTPAEVSIVVLSGTTAAQVTT 420

Db 421 tewwettarelpipepgdassimstesitgslgplldgtatrlrvkrqvpdcvlyry 480
Qy 421 TEWETTARELP IPEPGPDASSIMSTESITGSLGPLLDGTATRLRVKRQVP LDCVLYRY 480

Db 481 gsfvtdlvqgiesaeilqavpsgegdafeltvscggglpkacmeisepgcqpqagrl 540
Qy 481 GSFVTLDIVQGESAEILQAVP SEGDAFELTVSCGGGLPKACMEISEPGCQPQAGRL 540

Db 541 cqpvlspacqlvlhqlkqsgqtcylcnvsladtnslavvstqlimpqgaqlgvpliv 600
Qy 541 CQPVLSPACQLVLHQLKQSGTCYLCNVSLADTNSLAVVSTQLIMPQGAQLGVPLIV 600

Db 601 gillvlmavvlasliyrllmkqdfsvpqlphssahwlrilprifscpigenellsgqq 660
Qy 601 GILLVLMVVLASLIYRRLMKQDFSVPLPHSSSHWLRILPRIFCSCPIGENELLSGQQ 660

Db 661 v 661
Qy 661 V 661
```

```
RESULT 2
ENTRY A41234 #type complete
TITLE melanocyte-specific protein Pmel-17 precursor - human
ORGANISM #formal name Homo sapiens #common name man
DATE 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change
```

```
30-Sep-1993
ACCESSIONS A41234
REFERENCE A41234
#authors Kwon, B.S.; Chintamaneni, C.; Kozak, C.A.; Copeland, N.G.;
Gilbert, D.J.; Jenkins, N.; Barton, D.; Francke, U.;
Kobayashi, Y.; Kim, K.K.
#journal Proc. Natl. Acad. Sci. U.S.A. (1991) 88:9228-9232
#title A melanocyte-specific gene, Pmel 17, maps near the silver
coat color locus on mouse chromosome 10 and is in a
syntenic region on human chromosome 12.
#cross-references MIMD:92021023
#accession A41234
#status preliminary
#molecule type mRNA
##residues 1-668 ##label KWO
##cross-references GB:M77348
SUMMARY #length 668 #molecular-weight 70932 #checksum 6409

Query Match 97.8%; Score 4595; DB 11; Length 668;
Best Local Similarity 98.7%; Pred. No. 0.00e+00;
Matches 659; Conservative 0; Mismatches 2; Indels 7; Gaps 1;

Db 1 mdlvklrclhlhavigallavgatkvprnqdlgvsrqlrtkawnrqlypewteaqrldc 60
Qy 1 MDLVKRCLLHLAVIGALLAVGATKVP RNQDLGVS RQLRTKAWNRQLYP EWTEAQR LDC 60

Db 61 wrggqvskvndgptlliganasfsialnfpqskvlpdgqvlwvntlingsqvwgqp 120
Qy 61 WRGGQVSKVNDGPTLLIGANASFSIALNFPQSKVLPDGQVLWVNTLINGSQVWGQP 120

Db 121 vypeqtdacifpdgpcpsqgswqkrfvyvkwkgvqvlpgpvgaglsigtgramlg 180
Qy 121 VYPEQTDACIFPDGPCPSQGSWQKRFVYVWKWKGVLPGPVGAGLSIGTGRAMLG 180

Db 181 thtmvtyvhrgrsyrpplahsaeafitcdqvpfsveqlraldgnkhflrnpqltf 240
Qy 181 THTMVTVVHRGRSRYPP LAHSSAEFTITDQVPFSVSQLRALDGNKHFLRNPQLTF 240

Db 241 alqlhdpsglaeaadleytwdfgdsqgtllsralvvtthylepqpvtqavqlaaipits 300
Qy 241 ALQLHDP SGYLA EADLSYTWDFGDSGTLLSRA LVVTHYLEPQPVTQAVQLAAIPTS 300

Db 301 cgssvpvgttddghrptaeapnttaqvpvtevgtttqgaptaeapsgttsqvpttevis 360
Qy 301 CGSSVPVGTDDGHRPTAEAPNTTAQVPVTEVGTTPQAPTAEPSGTTSVQVPTTEVIS 360

Db 361 tapvqmpataestgmtpekvpvsevmgtllaemstpeatgmtpaevsiwvlgtaaqrvt 420
Qy 361 TAPVQMPATAESTGMTPEKVPVSEVMGTTLAEMSTPEATGMTPAEVSIVVLSGTTAAQVTT 420

Db 421 tewwettarelpipepgdassimstesitgslgplldgtatrlrvkrqvpdcvlyry 480
Qy 421 TEWETTARELP IPEPGPDASSIMSTESITGSLGPLLDGTATRLRVKRQVP LDCVLYRY 480

Db 481 gsfvtdlvqgiesaeilqavpsgegdafeltvscggglpkacmeisepgcqpqagrl 540
Qy 481 GSFVTLDIVQGESAEILQAVP SEGDAFELTVSCGGGLPKACMEISEPGCQPQAGRL 540

Db 541 cqpvlspacqlvlhqlkqsgqtcylcnvsladtnslavvstqlimpqgaqlgvpliv 600
Qy 541 CQPVLSPACQLVLHQLKQSGTCYLCNVSLADTNSLAVVSTQLIMPQGAQLGVPLIV 600

Db 601 gillvlmavvlasliyrllmkqdfsvpqlphssahwlrilprifscpigenellsgqq 660
Qy 601 GILLVLMVVLASLIYRRLMKQDFSVPLPHSSSHWLRILPRIFCSCPIGENELLSGQQ 660
```







	115-309	#region FESTV #region 9-residue repeats (D/E-T-T-T-A-S-T-T-A)\ #region 9-residue repeats (E-T-T-T-V-P-T-T-P)\ #region 4-residue repeats (E-T-T-T)\ #domain trefoil homology #label TRF3\ #product spasmodys II #status predicted #label SP2\ #domain trefoil homology #label TRF4\
	123-131, 231-239	
	132-230, 240-248	
	249-252, 276-299	
	305-342	
	351-400	
	353-393	
	23-49, 33-48, 43-60,, 74-100, 84-99, 94-111, 312-327, 322-339, 353-379, 363-378, 373-390	
	63	
SUMMARY	#length 400 #molecular-weight 42641 #checksum 1484	
Query Match	4.4%; Score 206; DB 5; Length 400;	
Best Local Similarity	30.1%; Pred. No. 1.41e-09;	
Matches	52; Conservative 34; Mismatches 80; Indels 7; Gaps 7;	
Db	122 pdttastaaetvptpettvtpptpettvtpptpettvtpptpettvtpptett 180	
Qy	283 PGVTAQVLQAALPLTSCGSPVP GTDCHRP TAEAPNTAGQPVTTEVGTTPGQAPT 342	
Db	181 v-pttpettvtpptpettvtp-tpttvtptpettvtpptpettvtp-ttpttaatt 237	
Qy	343 AEPSTGTSVQPTTEVISTAPVQMPTAESTGMTPEKYPVSEVMGTTLAEMSTEAT-GMT 401	
Db	238 taettvtpptpettpttpttdtpttlp-ptpettttettttettttet 289	
Qy	402 PAEVSVILSGTTAAQVTTTETWETTARELP IPEEGP-DASSIMSTESITGS 453	
RESULT	8	
ENTRY	A47282	#type complete
TITLE	calcium-binding protein calphotin - fruit fly ( <i>Drosophila melanogaster</i> )	
ORGANISM	#formal name <i>Drosophila melanogaster</i>	
DATE	21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 23-Mar-1995	
ACCESSIONS	A47282	
REFERENCE	A47282	
#authors	Martin, J.H.; Benzer, S.; Rudnicka, M.; Miller, C.A.	
#journal	Proc. Natl. Acad. Sci. U.S.A. (1993) 90:1531-1535	
#title	Calphotin: a <i>Drosophila</i> photoreceptor cell calcium-binding protein.	
#cross-references	MUID:93165729	
#accession	A47282	
#status	preliminary	
#molecule_type	nucleic acid	
#residues	1-865 ##label MAR	
#cross-references	NCBI:N124955; NCBI:P124956	
#experimental_source	photoreceptor cells	
#note	_sequence extracted from NCBI backbone	
SUMMARY	#length 865 #molecular-weight 84781 #checksum 2438	
Query Match	4.4%; Score 205; DB 10; Length 865;	
Best Local Similarity	23.7%; Pred. No. 1.83e-09;	
Matches	47; Conservative 63; Mismatches 79; Indels 9; Gaps 8;	
Db	114 vappvaatptpvqivpaapvi---atppvaacaptpaavtpvisvaspyvpantttvp 171	
Qy	270 ISRALVVHTLEFGVTAQQVLAQAALPLTSCGSPVPVGTGDGHRTAEAPNTAGC-VP 328	

```

Db   172 vaapvaavpaavpylalapavapavvpaetppppvaeipvatipecvaplipavs 231
      : ||| | : | : | : ||| : ||| : ||| | : |||
Qy   329 TTEWGTTPGQAPTAEPSCGTTVQVPTTEVTISTAPVQMP-TAE-STGMTPEKV-P-VSEVM 385

Db   232 vvatkplaaevpvvappatpetpvapaasphvsvapavetauva-pvsasetppvaa 290
      : ::::| ::::| || ::::::|| : ||| : ||| : | | :
Qy   386 GTTAAEMSTPENGMTPAEVSIVLSCGTTAAQTYYTENVEVTARELP-IEP-ECPDASS 443

Db   291 tltpatpetalapaavaes 308
      ::| : | : ||| : :
Qy   444 IMSTESITGSLGLDGT 461

RESULT          9
ENTRY           A39321    #type fragment
TITLE          mucin - rat (fragment)
ORGANISM       #formal name Rattus norvegicus #common name Norway rat
DATE           19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change
                  18-Jun-1993
ACCESSIONS     A39321
REFERENCE      A39321
#authors      Gum Jr., J.R.; Hicke, J.W.; Lagace, R.E.; Byrd, J.C.;
                Toribara, N.W.; Siddiki, B.; Fearney, F.J.; Lampert,
                D.T.A.; Kim, Y.S.
#journal      J. Biol. Chem. (1991) 266:22733-22738
#title        Molecular cloning of rat intestinal mucin. Lack of
                conservation between mammalian species.
#cross-references MUID:92042224
#accession    A39321
#status       preliminary
#molecule_type mRNA
#residues     1-447 #label GUM
#fcross-references GB:M76740
SUMMARY        #length 447 #checksum 7414

Query Match      4.3%; Score 200; DB 12; Length 447;
Best Local Similarity 29.0%; Pred.No. 6,63e-09;
Matches          49; Conservative 46; Mismatches 64; Indels 10; Gaps 10;

Db   139 vttdvtvvttvdvtttaevtttktvtsdpdattddtttpeattddgtvttpdvtttd 198
      ||| : ::| : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy   286 VTAQQVLQAIIPLTSCGS-SPVCTDGHRPTACANTTAGCVTTTEWGTPPGQATPE 344

Db   199 vt-tta-dvtttasvttfp-dvtttdpvttdp-dvtttagvtttdpdtattpdat-tp-q 252
      : ||| : || ::|| : | : | : | : | : | : | : | : | : | : | : | :
Qy   345 PSTGTSVQVPTTEVTISTAPVQMP-TAESTQMTPEKVPVSEWGTTLAEMSTPENGMTPE 404

Db   253 vtta-dvtttagvttagv-ttltpdvttt-pdvtttdpvttdvttasvttt 298
      : ||| : ||| |||| | ||| : ||| : | : | : | : | : | : | : | :
Qy   405 VSIWLSCGTTAAQTYYTENVEVTARELP-IEPEGPDASSIMSTESI TGS 453

RESULT          10
ENTRY           A53715    #type complete
TITLE          apomucin precursor - rat
ORGANISM       #formal name Rattus norvegicus #common name Norway rat
DATE           07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change
                  07-Oct-1994
ACCESSIONS     A53715
REFERENCE      A53715
#authors      Albone, E.F.; Hagen, F.K.; VanWyckhuysse, B.C.; Tabak, L.A.
#journal      J. Biol. Chem. (1994) 269:16845-16852
#title        Molecular cloning of a rat submandibular gland apomucin.
#accession    A53715
```

Aug 27 08:00

US-08-231-565A-27.rpr

11

```

##status      preliminary
##molecule_type  mRNA
##residues    1-322 ##label ALB
##cross-references  GB:U03407
KEYWORDS      glycoprotein; polymorphism; tandem repeat
FEATURE
1-22          #domain signal sequence #status predicted #label SIG\
23-322        #product apomucin #status predicted #label MAT
SUMMARY       #length 322 #molecular-weight 34975 #checksum 1086

Query Match      4.2%; Score 198; DB 12; Length 322;
Best Local Similarity 20.5%; Pred. No. 1.11e-08;
Matches 30; Conservative 48; Mismatches 68; Indels 0; Gaps 0;

Db 118 vtptqtkpttdsttpaptkpttdsttpaptkpttdsttpaptkpttdsttpaptn 177
Qy 306 VEGTDCGHRPTAEAPNTTAQVPTEVVGTPGQAPTAEPSGTSVQVPTTEVISTAPVQ 365

Db 178 kpttdsttpaptnkpttdsttpaptkpttdsttpaptkpttdsttpaptkpttdstt 237
Qy 366 MPTAESTGMPKVPVSEVWGTTTAEASTPEATGMPAEVSIWLSGTTAAQVTTTWEVE 425

Db 238 paptkpttdsttpaptkpttdstt 263
Qy 426 TTARELPPEPEGPDASSIMSTESIT 451

RESULT 11
ENTRY      S53363      #type fragment
TITLE      mucin JER58 - human (fragment)
ORGANISM   #formal name Homo sapiens #common name man
DATE       15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change
          01-Sep-1995
ACCESSIONS S53363
REFERENCE
#authors   Guyonnet-Duperrat, V.; Audie, J.P.; Debailleul, V.; Laine, A.;
          Buisson, M.P.; Galieque-Zouitina, S.; Pigny, P.; Degand,
          P.; Aubert, J.P.; Porchet, N.
#journal   Biochem. J. (1995) 305:211-219
#title     Characterization of the human mucin gene MUC5AC: a consensus
          cysteine-rich domain for 11p15 mucin genes?
#accession S53363
##status   preliminary
##molecule_type  mRNA
##residues 1-279 ##label GUY
##cross-references  EMBL:234278
##note     the nucleotide sequence is not given in this paper
SUMMARY    #length 279 #checksum 1256

Query Match      3.9%; Score 183; DB 11; Length 279;
Best Local Similarity 26.9%; Pred. No. 4.91e-07;
Matches 49; Conservative 43; Mismatches 84; Indels 6; Gaps 6;

Db 96 pspvptst-tea-pttrtsapissttatttstgpgtspvpttatt-sapttst 152
Qy 283 PCPVTAQVWLOAAIPLTSCSSPVGTTDCHR-PTAEAPNTTAGQVPTTEWGTTPQAP 341

Db 153 tsqg-gttsavppts-itvptstnsapissttsatttstgpgtspvptasts 210
Qy 342 TAEPSGTSVQVPTTEVISTAPVQMPAEAGTGPVKVPVSEVMGTTTAEASTPEATGWT 401

Db 211 asttstsgtspvpttstisvpttsttsttsttsttsttsttsttsttsttsttstt 270
Qy 402 PAEVSIVWLSGTTAAQVTTTWEVTTARELPPEPEGPDASSIMSTESITGSLGILDGT 461

```

Aug 27 08:00

US-08-231-565A-27.rpr

12

```

Db 271 st 272
Qy 462 AT 463

RESULT 12
ENTRY      A33532      #type fragment
TITLE      mucin SMUC-40 - human (fragment)
ORGANISM   #formal name Homo sapiens #common name man
DATE       29-Jan-1990 #sequence_revision 27-Jun-1994 #text_change
          27-Jun-1994
ACCESSIONS A33532
REFERENCE
#authors   Gum, J.R.; Byrd, J.C.; Hicks, J.W.; Toribara, N.W.; Lampport,
          D.T.A.; Kim, Y.S.
#journal   J. Biol. Chem. (1989) 264:6480-6487
#title     Molecular cloning of human intestinal mucin cDNAs. Sequence
          analysis and evidence for genetic polymorphism.
#cross-references  MUID:89197956
#accession A33532
##molecule_type  mRNA
##residues 1-167 ##label GUM
##cross-references  GB:M22404
KEYWORDS   glycoprotein
SUMMARY    #length 167 #checksum 7137

Query Match      3.7%; Score 175; DB 11; Length 167;
Best Local Similarity 27.4%; Pred. No. 3.56e-06;
Matches 43; Conservative 36; Mismatches 71; Indels 7; Gaps 7;

Db 9 stttmvtptptgtqtlptptttttt-tptptgtgtqtpstptsttttttptptpt 67
Qy 265 SSGTLISRALVWTH-T-YLEFGVTAQVWLOAAIPLTSCSSPVGTTDCHRPTEAPNTT 323

Db 68 gtqptttpttttttvtptptgtqtp-tpcp-itttttvt-ptptptg-tqtpptpi 123
Qy 324 AQQVPTTEVVGTPGQAPTAEPSGTSVQVPTTEVISTAPVQMPAEAGTGPVKVPVSE 383

Db 124 tttttvtptptgt-qtpstpttntttvtptptpt 159
Qy 384 VMGTTTAEASTPEATGMPAEVSIWLSGTTAAQVTT 420

RESULT 13
ENTRY      S55925      #type complete
TITLE      probable arabinogalactan protein precursor - tomato
ORGANISM   #formal name Lycopersicon esculentum #common name tomato
DATE       28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change
          03-Nov-1995
ACCESSIONS S55925
REFERENCE
#authors   Pogson, B.J.; Davies, C.
#journal   Plant Mol. Biol. (1995) 28:347-352
#title     Characterization of a cDNA encoding the protein moiety of a
          putative arabinogalactan protein from Lycopersicon
          esculentum.
#accession S55925
##status   preliminary
##molecule_type  mRNA
##residues 1-215 ##label POG
##cross-references  EMBL:247980
SUMMARY    #length 215 #molecular-weight 20812 #checksum 3703

```

```
Query Match          3.7%; Score 172; DB 9; Length 215;
Best Local Similarity 21.8%; Pred. No. 7.43e-06;
Matches 39; Conservative 50; Mismatches 85; Indels 5; Gaps 5;

Db      9 vsilciivasvtgtpaaavp-gakagtppaaaptkpktpapatapasapptaypva-p 66
      |: : : : : |: | | | | |: | | | | |: | | | | |: | | | |
Qy      286 VTAQWVLAALPITSCGSPVPGTTDCHRTAEAPNTTACQVPTTEVVGGTTPQQAETAP 345

Db      67 v-tapvtapttppva-apsvaspaapllkapasppvqspapapavattppavstppaaa 124
      |: | | | | | | | | | | |: | | | | | | |: | | | | | |
Qy      346 SGTTSVQVPTTEVISTAPVQWPTAEAGTPEKVPVSEVNGTTTIAENSTPEATGMTPAEV 405

Db      125 pvaapvasettpapaskdvkkgkghkhaepapdpdmsppapaeapypemdsdaa 183
      :: : : : : : : : : : : :: : : : : : : :: : : : : : :
Qy      406 SIVLSGTTAAQVTTTETWETTARELP IPEPEGDASSINST-ESITGSLGLDGTAT 463

RESULT 14
ENTRY   B33532      #type fragment
TITLE   mucin SMUC-41 - human (fragment)
ALTERNATE_NAMES mucin HAM-1
ORGANISM #formal name Homo sapiens #common name man
DATE     29-Jan-1990 #sequence_revision 27-Jun-1994 #text_change
          23-Mar-1995

ACCESSIONS B33532; A61257
REFERENCE   A33532
AUTHORS     Gum, J.R.; Byrd, J.C.; Hicks, J.W.; Toribara, N.W.; Lampport,
            D.T.A.; Kim, Y.S.
JOURNAL     J. Biol. Chem. (1989) 264:6480-6487
TITLE       Molecular cloning of human intestinal mucin cDNAs. Sequence
            analysis and evidence for genetic polymorphism.
CROSS-REFERENCES MUID:89197956
ACCESSION   B33532
MOLECULE_TYPE mRNA
RESIDUES    1-278 ##label GUM
CROSS-REFERENCES GB:M22405
EXPERIMENTAL_SOURCE Intestine
REFERENCE   A61257
AUTHORS     Jany, B.H.; Gallup, M.W.; Yan, P.S.; Gum, J.R.; Kim, Y.S.;
            Basbaum, C.B.
JOURNAL     J. Clin. Invest. (1991) 87:77-82
TITLE       Human bronchus and intestine express the same mucin gene.
ACCESSION   A61257
STATUS      not compared with conceptual translation
MOLECULE_TYPE mRNA
RESIDUES    'n', 10-33, 'TTS', 37-39 ##label JAN
EXPERIMENTAL_SOURCE bronchus
COMMENT      Southern blot analysis suggests that a polymorphic gene encodes
            both intestinal and bronchial mucin.

GENETICS
GENE         GDB:MUC2
MAP_POSITION l1p15.5
KEYWORDS     glycoprotein; polymorphism
SUMMARY      #length 278 #checksum 4056

Query Match          3.7%; Score 176; DB 11; Length 278;
Best Local Similarity 28.0%; Pred. No. 2.79e-06;
Matches 46; Conservative 41; Mismatches 69; Indels 8; Gaps 6;

Db      6 aptptpistttvtptptgt-gtpttpttittttvtptptgt-qtpttvlitttt 62
      :| : | : | : | | | | | | :| : | : | : | | | | :| : | : | : |
Qy      304 SVPCTDGHRTAEAPNTTACQVPTTEVVGGTTPQQAETAPESCTTSVQVPTTEVISTAP 363

Db      63 tmtptptskettvt-pittttttvtatptptgt-qtptmipistttttvtptpttgs 120
```

```
Qy      364 VOMPTAESTGMTPEKVPVSEVNGTTTIAENSTPEATGMTPAEVSIVLSGTTAAQVTTTEM 423
      |: | : | : | : | : | | : | | : | | : | | : | | : | |
Db      121 tgppthtstapiaeltnpppesstptqststaspltesttll 164
      | : | | | : : : : | | : : | : | : | : | : |
Qy      424 VET-T-ARELP IPEPEGDASSINST-ESITGSLGLDGTATL 464

RESULT 15
ENTRY   VGBEX1      #type complete
TITLE   glycoprotein X precursor - equine herpesvirus 1 (strain Ab4p)
ORGANISM #formal name equine herpesvirus 1
NOTE     host Equus caballus (domestic horse)
DATE     30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change
          17-Feb-1995

ACCESSIONS H36802
REFERENCE   A36805
AUTHORS     Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
SUBMISSION submitted to GenBank, March 1992
DESCRIPTION The DNA sequence of equine herpesvirus-1.
ACCESSION   H36802
MOLECULE_TYPE DNA
RESIDUES    1-797 ##label TEL
CROSS-REFERENCES GB:M86664
REFERENCE   A41831
AUTHORS     Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
JOURNAL     Virology (1992) 189:304-316
TITLE       The DNA sequence of equine herpesvirus-1.
CROSS-REFERENCES MUID:92295566
CONTENTS    annotation; possible protein-coding frames
NOTE        neither amino acid nor nucleotide sequence is given

GENETICS
GENE         71
CLASSIFICATION #superfamily equine herpesvirus glycoprotein X; equine
            herpesvirus 1 glycoprotein homology
            glycoprotein; transmembrane protein
KEYWORDS     1-22      #domain signal sequence #status predicted #label SIG\
            23-797     #product glycoprotein X #status predicted #label MAT\
            23-465     #region serine/threonine-rich\
            489-797     #domain equine herpesvirus 1 glycoprotein homology
            #label EH6\
            766-790     #domain transmembrane #status predicted #label TMN\
            590         #binding_site carbohydrate (Asn) (covalent) #status
            predicted
SUMMARY      #length 797 #molecular-weight 80342 #checksum 6547

Query Match          3.6%; Score 169; DB 3; Length 797;
Best Local Similarity 17.7%; Pred. No. 1.54e-05;
Matches 32; Conservative 64; Mismatches 84; Indels 1; Gaps 1;

Db      51 septtspttsspsptstststsgstsgstsgstsgstsgstsgstsgstsgstsgst 110
      ::| | : : | : | | | | : | | | : | | | : | | | : | | |
Qy      283 PGVTAQWVLAALPITSCGSPVPGTTDCHRTAEAPNTTACQVPTTEVVGGTTPQQAET 342

Db      111 tpttastttpttttaapttaattavttaastsaaettatat-atstptttttpttttt 169
      ::| | : : | | | | | | : | | | : : : : | : : : | | : | |
Qy      343 AEPSTTSVQVPTTEVISTAPVQWPTAEAGTPEKVPVSEVNGTTTIAENSTPEATGMT 402

Db      170 atttpttasttttdtttaatttaatttaatttaatttaatttaatttaatttaatt 229
      | : : : : | | : : | : | : | : | : | : | : | : : : :
Qy      403 AEVSIVLSGTTAAQVTTTETWETTARELP IPEPEGDASSINSTESITGSLGLDGT 462

Db      230 t 230
```

Aug 27 08:00

US-08-231-565A-27.rpt

15

Oy 463 T 463

Search completed: Tue Aug 27 08:12:37 1996  
Job time : 102 secs.

\*\*\*\*\*

WVLPSELA (TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Aug 27 08:09:36 1996; MasPar time 16.82 Seconds

Tabular output not generated.  
683.05% Million cell updates/sec

Title: &gt;US-08-231-565A-27

Description: (1-661) from US08231565A.pep

Perfect Score: 4698

Sequence: 1 MDVLKRLGLHLAVIGALLA.....RIFCSCPGENSPILLSGQGV 661

Scoring table: PAM 150

Gap 11

Searched: 49340 seqs, 17385503 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: swiss-prot32

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9

Statistics: Mean 52.781; Variance 121.768; scale 0.433

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	4586	97.6	668	6	PM17_HUMAN	PMEL 17 PROTEIN PRECU	0.00e+00
2	206	4.4	400	5	MU1_XENLA	INTEGUMENTARY MUCIN A	6.74e-13
3	205	4.4	865	2	CPN_DROME	CALPHOTIN.	9.36e-13
4	169	3.6	797	8	VGLX_HSVB	GLYCOPROTEIN X PRECUR	8.77e-08
5	169	3.6	907	8	VGP3_EBV	ENVELOPE GLYCOPROTEIN	8.77e-08
6	167	3.6	1367	1	AMVH_YEAST	GLUCAMYLASE SL/S2 PR	1.62e-07
7	164	3.5	600	7	SP96_DICDI	SPORE COAT PROTEIN SP	4.02e-07
8	163	3.5	2274	5	MUC2_HUMAN	MUCIN 2 (INTESTINAL M	5.44e-07
9	156	3.3	662	5	MU1_XENLA	INTEGUMENTARY MUCIN C	4.41e-06
10	148	3.2	670	8	VG50_HSV1	HYPOTHETICAL GENE 50	4.58e-05
11	148	3.2	671	1	ALYS_ENTFA	AUTOLYSIN (EC 3.5.1.2	4.58e-05
12	149	3.2	739	6	OCT1_CHICK	OCTAMER-BINDING TRANS	3.43e-05
13	149	3.2	743	6	OCT1_HUMAN	OCTAMER-BINDING TRANS	3.43e-05

14	144	3.1	474	9	VTP3_TTVIV	VIRAL PROTEIN TPX.	1.44e-04
15	147	3.1	742	1	BAL_HUMAN	BILE-SALT-ACTIVATED L	6.11e-05
16	145	3.1	1251	9	Y0U3_CAEEL	HYPOTHETICAL 133.5 KD	1.09e-04
17	141	3.0	224	9	Y091_NPVAC	HYPOTHETICAL 24.1 KD	3.38e-04
18	141	3.0	542	6	PSGL_ONCMY	APOLYSIALOGLYCOPROT	3.38e-04
19	142	3.0	610	5	MUC4_HUMAN	MUCIN 4 (TRACHEOBRONC	2.55e-04
20	141	3.0	796	9	YS8A_CAEEL	HYPOTHETICAL 84.3 KD	3.38e-04
21	140	3.0	1025	7	SLAP_CAUCR	S-LAYER PROTEIN (PARA	4.48e-04
22	138	2.9	398	5	MU1_XENLA	INTEGUMENTARY MUCIN B	7.84e-04
23	137	2.9	567	2	CH13_CANAL	CHITINASE 3 PRECURSOR	1.04e-03
24	136	2.9	605	9	YHC8_YEAST	HYPOTHETICAL 63.8 KD	1.37e-03
25	137	2.9	699	8	VGLG_HSV2	GLYCOPROTEIN G.	1.04e-03
26	134	2.9	705	4	GUN6_DICDI	ENDOGUCANASE PRECURS	2.37e-03
27	138	2.9	725	1	AGAL_YEAST	A-AGGLUTININ ATTACHEME	7.84e-04
28	132	2.8	173	4	HA34_BRELK	HAM34 PROTEIN.	4.09e-03
29	132	2.8	259	5	MSP8_EIMAC	MEROZOITE SURFACE PRO	4.09e-03
30	132	2.8	353	7	SUC3_RAT	SYNDECAN-3 (N-SYNDECA	4.09e-03
31	130	2.8	361	6	PIR3_YEAST	PIR3 PROTEIN PRECURSOR	7.04e-03
32	130	2.8	573	1	C114_MOUSE	CELL SURFACE ANTIGEN	7.04e-03
33	131	2.8	630	5	MU1_MOUSE	MUCIN 1 PRECURSOR (PO	5.37e-03
34	128	2.7	400	5	LEUK_HUMAN	LEUKOSIALIN PRECURSOR	1.20e-02
35	129	2.7	676	4	ICP0_HSVBJ	TRANS-ACTING TRANSRI	9.21e-03
36	126	2.7	1609	9	YCA9_YEAST	HYPOTHETICAL 166.0 KD	2.05e-02
37	129	2.7	1849	4	IGA4_HAETN	IMMUNOGLOBULIN A1 PRO	9.21e-03
38	127	2.7	1854	2	CIPA_CLOTM	CELLULOSOMAL SCAFFOLD	1.57e-02
39	127	2.7	2441	2	CBP_MOUSE	CREB-BINDING PROTEIN.	1.57e-02
40	123	2.6	338	7	SEDI_YEAST	SEDI PROTEIN PRECURSOR	4.51e-02
41	123	2.6	439	9	XP2_XENLA	SKIN SECRETORY PROTEI	4.51e-02
42	124	2.6	713	8	TUP1_YEAST	GLUCOSE REPRESSION RE	3.47e-02
43	124	2.6	767	1	AMVH_SACDI	GLUCOAMYLASE S1 PRECU	3.47e-02
44	123	2.6	1150	1	APMU_PIG	APOMUCIN (MUCIN CORE	4.51e-02
45	123	2.6	1162	8	TCNA_TRYCR	SIALIDASE (EC 3.2.1.1	4.51e-02

## ALIGNMENTS

RESULT	1	PM17_HUMAN	STANDARD;	PRT;	668 AA.
ID	AC	P40967;			
DT	01-FEB-1995	(REL. 31, CREATED)			
DT	01-FEB-1995	(REL. 31, LAST SEQUENCE UPDATE)			
DT	01-FEB-1995	(REL. 31, LAST ANNOTATION UPDATE)			
DE	PMEL 17 PROTEIN PRECURSOR.				
GN	PMEL17.				
OS	HOMO SAPIENS (HUMAN).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; PRIMATES.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 92021023.				
RA	KNON B.S., CHINTAMANENI C., KOZAK C.A., COPELAND N.G.,				
RA	GILBERT D.J., JENKINS N., BARTON D., FRANCKE U., KOBAYASHI Y.,				
RA	KIM K.-K.;				
RL	PROC. NATL. ACAD. SCI. U.S.A. 88:9228-9232(1991).				
CC	-/- FUNCTION: COULD BE A MELANOGENIC ENZYME.				
CC	-/- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN MELANOCYTES.				
CC	-/- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).				
DR	EMBL; M77348; M77348.				
DR	MIM; 155550; 11TH EDITION.				
KW	TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL; MELANIN BIOSYNTHESIS; REPEAT.				
FT	SIGNAL	1	23	POTENTIAL.	
FT	CHAIN	24	668	PMEL 17 PROTEIN.	
FT	TRANSMEM	575	595	POTENTIAL.	
FT	TRANSMEM	603	623	POTENTIAL.	





QY 343 AEPGTTSVQVPTTEVISTAPVQMPATAESTGMTPEKVPVSEVMGTTLAEMSTPEAT-GMT 401

Db	238	taettvpttpettteptttpttdtgpptlp-ptpetttttttttttttt	289
		:   :         :   :           :   :     :	
QY	402	PAEVSIVLSCGTAQVTTTQVTTTETARELPPEEGP-DASSIMSTES	453

RESULT	3	
ID	CPN DROME	STANDARD;
		PRT; 865 AA.

AC	Q02910:	
DT	01-OCT-1993	(REL. 27, CREATED)
DT	01-OCT-1993	(REL. 27, LAST SEQUENCE UPDATE)
DT	01-FEB-1994	(REL. 28, LAST ANNOTATION UPDATE)
DE	CALPOTIN.	
OS	CPN OR CAP.	
OC	DROSOPHILA MELANOGASTER (FRUIT FLY).	
OC	EUKARYOTA: METAZOA: ARTHROPODA: INSECTA: DIPTERA:	

[1]	SEQUENCE FROM N.A.
RP	RC
RP	STRAIN=CANTON-S;
RX	MEDLINE; 93165729.
RA	MARTIN J.H., BENZER S., RUDNICKA M., MILLER C.A.;
RL	PROC. NATL. ACAD. SCI. U.S.A. 90:1531-1535(1993).

[2] SEQUENCE FROM N. A.  
RP STRAIN=CANTON-S;  
RX MEDLINE; 93165730.  
RA BALLINGER D.G., XUE N., HARSHMAN K.D.;  
RL PROC. NATL. ACAD. SCI. U.S.A. 90:1536-1540(1993).  
CC -1- FUNCTION: MIGHT FUNCTION AS A CALCIUM-SEQUESTERING "SPONGE" TO  
CC REGULATE THE AMOUNT OF FREE CYTOPLASMIC CALCIUM. IT BINDS 0.3 MOL  
CC OF CA+2 PER MOL OF PROTEIN.

CC	-1-	SUBUNIT: HOMODIMER (PROBABLE).
CC	-1-	SUBCELLULAR LOCATION: CYTOPLASMIC; HYPODENSE COMPARTMENT.
CC	-1-	DEVELOPMENTAL STAGE: EXPRESSED EARLY IN PHOTORECEPTOR CELL DEVELOPMENT.
CC	-1-	TISSUE SPECIFICITY: SOMA AND AXONS OF PHOTORECEPTOR CELLS
CC		COMPOUND EYES AND OCELLI

DR EMBL; L02111; L02111.  
DR EMBL; L05080; L05080.  
DR PIR; A47282; A47282.  
DR FLYBASE; FBGN0010218; CPN.  
KW CALCIUM-BINDING.

FT	CONFLICT	36	36	A -> AVAPAWA (IN REF. 2).
FT	CONFLICT	43	43	I -> T (IN REF. 2).
FT	CONFLICT	64	64	I -> V (IN REF. 2).
FT	CONFLICT	76	76	T -> A (IN REF. 2).
FT	CONFLICT	100	100	P -> PP (IN REF. 2).
FT	CONFLICT	126	127	VQ -> AP (IN REF. 2).
FT	CONFLICT	154	154	I -> V (IN REF. 2).
FT	CONFLICT	160	160	S -> T (IN REF. 2).
FT	CONFLICT	534	534	A -> E (IN REF. 2).
FT	CONFLICT	699	699	I -> T (IN REF. 2).
FT	CONFLICT	703	703	V -> L (IN REF. 2).
FT	CONFLICT	721	721	D -> E (IN REF. 2).
SQ	SEQUENCE	865 AA;	84781 MW;	4162550 CN;

Query Match 4.4%; Score 205; DB 2; Length 865;  
Best Local Similarity 23.7%; Pred. No. 9,36e-13;  
Matches 47; Conservative 53; Mismatches 79; Indels 9; Gaps 8;

[illegible]

Db	172	vaapvaavpaavpwvpylapavapavpvvaetpappvaaipvatipectvaplpevs	231
		:   :   :   :   :   :   :   :   :   :   :   :   :   :	
Qy	329	TTEWGTGQAPTAEPSGTTSGVPTTEVISTAPVQMPTAE-STGMTPEKV-P-VSEVM	385
		:   :   :   :   :   :   :   :   :   :   :   :   :   :	
Db	232	vwatkplaaepvvwvappatetpvvpaasaaphvsvapavetaavva-pvsaetepvva	290
Qy	386	GTTTIAEMSTPEATGMTPEVSI-VVLSGTAAQVTTTTEVETARELP-IEP-EGPDASS	443
		:   :   :   :   :   :   :   :   :   :   :   :   :   :	
Db	291	tltpatetpalapvvaes	308
Qy	444	IMSTISITGSLGLDGT	461

RESULT	4	
ID	VGLX HSVEB	STANDARD;
		PRT; 797 AA.

AC	P28958;
AD	01-DEC-1992 (REL. 24, CREATED)
DT	01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT	01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
DT	01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
DE	GLYCOPROTEIN X PRECURSOR.
GN	71.
OS	EQUINE HERPESVIRUS TYPE 1 (STRAIN AB4P) (EHV-1).
OC	VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; ALPHAHERPESVIRINAE.
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE; 92295566.
RL	RA TELFORD E.A.R., WATSON M.S., MCBRIDE K., DAIVSON A.J.; VIROLOGY 189:304-316(1992).

DR	EMBL; M86664; M86664.
DR	PIR; H36802; VGEX1.
KW	GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL.
FT	1 22
FT	SIGNAL
FT	23 797
FT	CHAIN
FT	23 465
FT	DOMAIN
FT	766 790
FT	TRANSMEM
FT	POTENTIAL.
FT	CARBOHYD 590 590
FT	POTENTIAL.
SQ	SEQUENCE 797 AA; 80342 MW; 3610380 CN;

Query Match 3.6%; Score 169; DB 8; Length 797;  
Best Local Similarity 17.7%; Pred. No. 8.77e-08;  
Matches 32; Conservative 64; Mismatches 84; Indels 1; Gaps 1;

**D**b      51 ssppttesspptethsspestsqsstassapstastett 110  
:: :: :: :: :: :: :: :: :: :: :: :: :: :: ::  
**Qy**    293 pcvrtvqvllqaipitscgsspvpcttdgrprtaepntagovttewgtwtpgaqt 342

[illegible][illegible]

Qy 403 AEVSIWLSGTTAAQVTTTEWVETTARELP IPEPECPDASSIMSTESTIGSLP LIDGTA 462

Db 230 t 230

Qy 463 T 463

RESULT	5	
ID	VGP3	EBV
		STANDARD;
		PRT; 907 AA.

AC P03200;  
DT 21-JUL-1986 (REL. 01, CREATED)

US-08-231-565-A-27 JSP  
AUG 27 07:58

US-08-231-565A-27.rsp

421 fsekapeattstplnttgfadpntttglpesthvptnltapastgvtadvstadsptpag 480  
l::  
262 FCGSSGTLIRALVTHYLEPCVTA-QVVLLQAAILITSCGS-PVPGTTDGHRTAEFA 319  
::  
481 ttgasvpvtpspwdngteeskacdm-tsstsvtttpnatsptpavtttpnatspt 539  
::  
320 PNMTAGQV-PTTEV-VGTTPGQAPTAEPSTSVQVPTEVISATP-VQMPTAESICWTP 376  
::  
540 avtptpnatsptlgktsptsavtttpnatsptlgktsptsavtttpnatsptlgkts 599  
l::  
377 E-KPVSEVMGVNLTIAEMSTPEATGMTPAETSIVVLISGTTAAQTVTTFENVEVETARELPIPE 435  
::  
600 ptsavtttpnatsptgvgtetgsqanathl 630  
l::  
436 PEGPDASSI-MSTESITSGSLGLDCTA-TL 464  
:: ::

### RESULT 6

ID	ANYH YEAST	STANDARD;	PRT; 1367 AA.
AC	P08640; P08068;		
DR	01-AUG-1988 (REL. 08, CREATED)		
DT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)		
DDT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)		
DE	GLUCOAMYLASE S1/S2 PRECURSOR (EC 3.2.1.3) (GLUCAN 1,4-ALPHA-		
DE	GLUCOSIDASE) (1,4-ALPHA-D-GLUCAN GLUCOHYDROLASE).		
GN	STA1 OR STA2 OR WAL5 OR YIR019C.		
OC	SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).		
CC	EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RP	STRAIN=S288C / AB9712;		
RR	BARELL B.G., BADCOCK K., BANKIER A.T., BOWMAN S., BROWN D.,		
RRA	CHURCHER C.M., CONNOR R., COPSEY T., DEAR S., DEVLIN K., FRASER A.,		
RRL	GENTLES S., HAMLYN N., HORSNELL T.S., HUNT S., JAGELS K., JONES M.,		
RRA	LOUIS E., LYE G., MOULE S., MOULE T., ODELL C., PEARSON D.,		
RRA	RAJANDREWM M.A., RILES L., ROWLEY N., SKELTON J., SMITH V.,		
RRA	WALSH S.V., WHITEHEAD S.;		
RRL	SUBMITTED (DEC-1994) TO EMBL/GENBANK/DDJB DATA BANKS.		
RN	[2]		
RC	SEQUENCE OF 1-242 AND 762-1331 FROM N.A.		
RR	MEDLINE: 87194600.		
RRL	YAMASHITA I., NAKAMURA M., FUKUI S.;		
RRA	J. BACTERIOL. 169:2142-2149(1987).		
RN	[3]		
RC	SEQUENCE OF 1-31 FROM N.A.		
RR	STRAIN=SPX101-1C;		
RRL	MEDLINE: 89031230.		
RRA	PARGO J.M., INEZ E., ZALACAIN M., CLAROS M.G., JIMENEZ A.;		
RRL	FEEBS LETT. 239:179-184(1988).		
CCC	-1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL 1,4-LINKED ALPHA-D-		
CCC	GLUCOSE RESIDUES SUCCESSIVELY FROM NON-REDUCING ENDS OF THE CHAINS		
CCC	WITH RELEASE OF BETA-D-GULOSE.		
CCC	-1- SIMILARITY: BELONGS TO FAMILY 15 OF GLYCOSYL HYDROLASES.		
DR	EMBL; Z47047; Z47047.		
DR	EMBL; M16164; M16164.		
DR	EMBL; M16165; M16165.		
DR	EMBL; X13857; X13857.		
DR	PIR; B26877; B26877.		
DR	PIR; A26877; A26877.		
DR	PIR; S48478; S48478.		
DR	LISTA; SC01218; STAI.		
DR	SGD; L0002104; STAI.		
DR	PROSITE; PS00820; GLUCOAMYLASE.		

Query Match 3.6%; Score 169; DB 8; Length 907;  
Best Local Similarity 23.2%; Pred. No. 8.7/e-08;  
Matches 49; Conservative 56; Mismatches 97; Indels 9; Gaps 9;





FT	DISULFID	573	599	BY SIMILARITY.
FT	DISULFID	583	598	BY SIMILARITY.
FT	DISULFID	593	610	BY SIMILARITY.
FT	DISULFID	621	647	BY SIMILARITY.
FT	DISULFID	631	646	BY SIMILARITY.
FT	DISULFID	641	658	BY SIMILARITY.
FT	VARIANT	276	276	K → E.
FT	VARIANT	354	354	C → R.
FT	VARIANT	415	415	T → A.
SQ	SEQUENCE	662 AA;	67774 MW;	2659094 CN;

```

Query Match      3.3%; Score 156; DB 5; Length 662;
Best Local Similarity 28.3%; Pred. No. 4,41e-06;
Matches 34; Conservative 24; Mismatches 61; Indels 1; Gaps 1;

```

Db 454 tpt-ttttkatttttptttttptttttkattttptttttptttttttttt 512  
: | : | | | : | | : | | : | | : | : | | |  
Qv 362 APVOMPTAESTGHTPEKVPVSEWVGTTLAEMSTPEATGHTPAEVSIVVLGGTAAOVTTT 421

RESULT	10
ID	VC50_HSVI1
AC	Q00130;
DT	01-DEC-1992 (REL. 24, CREATED)
DT	01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT	01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
DE	HYPOTHETICAL GENE 50 PROTEIN.

30. Ictalurid herpesvirus 1 (channel catfish virus) (CCV).  
OS VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; ALPHAPERESVIRINAE.  
RN [1]

RP	SEQUENCE FROM N.A.
RC	STRAIN=AUBURN 1;
RX	MEDLINE; 92087490.
RA	DAVISON A.J.;
RL	VIROLOGY 186:9-14 (1992).
DR	EMBL; M75136; M75136.
DR	PIR; F36791; F36791.
DR	HYPOTHETICAL PROTEIN.
SK	SEQUENCE 670 AA; 64174 MW;
QW	SEQUENCE 670 AA; 2495347 CN;

Query Match 3.2%; Score 148; DB 8; Length 670;  
Best Local Similarity 22.0%; Pred. No. 4.58e-05;  
Matches 39; Conservative 48; Mismatches 80; Indels 10; Gaps 10;

83 psiahttskmielgtprptipgvtt-ipntdap-vdpgsvhttarvvtidittkqpt 140  
| | : : : : | : : | : : | : : | : :  
285 PVTAQVVLQAALPLTSCGSSPVFTGDGHRTAEAPNTTAGQVPTE-VW-GTTPGQAPT 342

Db 141 tpatpagandantittatpagandantittatpagandantittatpagandtattt 200  
Qy 343 A-E-PSGTT-SVQVPT-TEVTSTAPQMPTAESTCMT-PEKVPSEVMCTTIAEMSTPEA 397

Db 201 tpagandtavtttptapagandtandvtttppampagandtangtawtttppampag 257  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :  
Qy 398 TGMTPAEVSIVVLGGTTA-AQVTTTETWVETTARELPPEPEGDASSINTESTITGS 453

RESULT	11	
ID	ALYS ENTFFA	STANDARD; PRT; 671 AA.

AC	P37710;	
AD	01-OCT-1994 (REL. 30, CREATED)	
DT	01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)	
DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)	
DE	AUTOLYSIN (EC 3.5.1.28) (N-ACETYLMURAMYL-L-ALANINE AMIDASE) .	
OS	ENTEROCOCCUS FAECALIS (STREPTOCOCCUS FAECALIS) .	
OC	PROKARYOTA; FIRMICUTES; COCCI; STREPTOCOCCAEAE.	
RN	(1)	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE; 91358349.	
RA	BELIVEAU C., POTVIN C., TRUDEL J., ASSELIN A., BELLEWARE G.;	
RL	J. BACTERIOL. 173:5619-5623(1991).	
CC	-!- FUNCTION: HYDROLYZES THE CELL WALL OF E.FAECALIS AND	
CC	M.LYSODEIKTICUS.	
CC	-!- CATALYTIC ACTIVITY: HYDROLYSES THE LINK BETWEEN N-ACETYLMURAMOYL	
CC	RESIDUES AND L-AMINO ACID RESIDUES IN CERTAIN BACTERIAL CELL-WALL	
CC	GLYCOPEPTIDES.	
CC	EMBL; M58002; M58002.	
DR	PIR; A38109; A38109.	
KW	HYDROLASE; CELL WALL; REPEAT.	
FT	FT DOMAIN 348 613 4 X 67 AA APPROXIMATE TANDEM REPEATS.	
FT	REPEAT 348 414 1.	
FT	REPEAT 415 482 2.	
FT	REPEAT 483 550 3.	
FT	REPEAT 551 616 4.	
SO	SEQUENCE 671 AA; 70442 MW; 2419279 CN;	

Query Match 3.2%; Score 148; DB 1; Length 671;  
Best Local Similarity 28.2%;  
Pred. No. 4.58e-05;  
Matches 55; Conservative 49; Mismatches 76; Indels 15; Gaps 13;

Db 27 sttlfssali vsvsgtgvallpvtaeateeqptn-aevaqa pttetglvetpttettpgi 85  
| ||| | ||| : | |||| : : : : | ||| :  
Qv 266 SGTILSRALVWTH--TVLEPGFVTVQVQLAAIPLTSCGSSPVGTTDGHRTAEAPNTT 323

```

Db      86  teq-ptdssttestessketptpstreqptvd--sttpvesgttdss-va-eitpva 140
      : |||: ||: :: : | : | : | : |||: ||: :: : | : ||:
Qy     324  AGQVPTTEVWGTPGQAPTA-EPSGTTSVQVPTTEVISTAPVQMPSTAESTGMTPEKVPVS 382

```

Db 141 psttesaapavtpdddevkvppearvasaqtfalsptq-spsefielarcagpiaand 199  
Qv 383 -EVNGTTIAENSTPEATGMP-AEVSIV-VLSGTAAQTWTTTETWTAR-ELPIEPEG 438

Db 200 *lyas-vmmaqvaives* 213  
 || :| ::| |  
 Qv 439 *PDASSIMSTESITGS* 453

RESULT	12	OTC1 CHICK	STANDARD;	PRT;	739 AA.
ID	AC	P15143;			
DT	DT	01-APR-1990	(REL. 14, CREATED)		
DT	DT	01-APR-1990	(REL. 14, LAST SEQUENCE UPDATE)		
DT	DT	01-JUL-1993	(REL. 26, LAST ANNOTATION UPDATE)		
DE	DE	OCTAMER-BINDING TRANSCRIPTION FACTOR 1 (OTF-1)			(NF-A1).
GN	GN	OTF1 OR OCT1.			

OC GALLOS GALLINAE (CHICKENS):  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;  
OC CALLIFORMES.  
[1]  
RN  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 90138945.  
RA PETRYNIAK B., POSTEMA C.E., MCCORMACK W.T.,  
RA THOMPSON C.B.;

RESULT	14	ID	VT3	TTV1	STANDARD;	PRT;	474	AA.
AC	P19275;							
DT	01-NOV-1990	(REL. 16,	CREATED)					
DT	01-NOV-1990	(REL. 16,	LAST SEQUENCE UPDATE)					
DT	01-FEB-1991	(REL. 17,	LAST ANNOTATION UPDATE)					
DE	VIRAL PROTEIN	TPX.						
OS	THERMOPROTEUS TENAX	VIRUS 1	(STRAIN VT3)	(TTV1).				
OC	VIRIDAE;	DS-DNA	ENVELOPED	VIRUSES;	LIPOTHRIXVIRIDAE.			
RN	[1]							
RP	SEQUENCE FROM	N.A.						
RX	MEDLINE;	90245666.						
RA	NEUMANN H.,	ZILLIG W.;						
RL	NUCLEIC ACIDS	RES. 18:	2171-2171	(1990).				
DR	EMBL;	X14717;	X14717.					

RESULT	13	ID	OCT1 HUMAN	STANDARD;	PRT;	743 AA.
		AC	PI4859;			
		DT	01-APR-1990 (REL. 14, CREATED)			
		DT	01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)			
		DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)			
		DE	OCTAMER-BINDING TRANSCRIPTION FACTOR 1 (OTF-1)			(NF-A1) .
		GN	OTF1 OR OCT1.			
		OS	HOMO SAPIENS (HUMAN) .			
		OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
		OC	EUTHERATA; PRIMATES.			

[1] RP SEQUENCE FROM N.A.  
RX MEDLINE; 8910793.  
RA STURM R.A., DAS G., HERR W.;  
RRL GENES DEV. 2:1582-1599(1988).  
[2] RN  
RP STRUCTURE BY NMR OF 284-359.  
RX MEDLINE; 93241300.  
RA DEKKER N., COX M., BOELEN S R.,  
RA KAPTEIN R.;  
RL NATURE 362:852-855(1993).  
[3] RN  
RP STRUCTURE BY NMR OF 284-354.  
RX MEDLINE; 93214991.





FT REPEAT 724 734  
SQ SEQUENCE 742 AA; 78345 MW; 3107924 CN;  
Query Match 3.1%; Score 147; DB 1; Length 742;  
Best Local Similarity 24.6%; Pred. No. 6.11e-05;  
Matches 42; Conservative 50; Mismatches 67; Indels 12; Gaps 12;  
Db 567 eatpvpptgdsdapvptgdsqgappvptgdsqgappvptgdsqgappv 626  
Qy 282 EPGVTHAQVWLQAA-IPLTS-CGSSPVPCTD-GHRPTAEAPNTTACQVPTTEVVGTTTG 338  
Db 627 -pptgd-sqappvp-ptqds-qappvp-ptqds-q-appvptgdagppvpvptgdsqap 679  
Qy 339 QAPTAEPSGTSVQVPTTEVISTAPVQMPTEASTGMPTEKVPVSEVMGTTLAEMSTPEAT 398  
Db 680 pvpptgdsqappvptgdsdapvptgdsqg-appvptgdsdapvptgd 729  
Qy 399 GMTPA-EVSIVVLGTTAAQVTTTEWVETTARELPPEPEGPDASSIMSTE 448

Search completed: Tue Aug 27 08:10:37 1996  
Job time : 61 secs.



[illegible]

Db	382	gtgtatccccaggaataactgacgattgcttgcattctccctgatgttgagcttggccattc	441
Qy	361	gtttatccnccargacnagcaygagctgctatthtttccnagcaygngcnctgycnns	420
Db	442	ggctcttggtctcagaagaagcttctgtttactgtctggaagcattggggccaatactgg	501
Qy	421	ggnsntgsgncsraaragwnsnttctgttctgttggaaracntgggncartaytgg	480
Db	502	caagttctadggggcccagtgctgggctgagcattgggacaggcaggcgaatgctgggc	561
Qy	481	carttcttggcngcngcngtwnsgcnytnwsnathgncacngnmgcnatgctyngcn	540
Db	562	acacacacattggaagtgaetgctaccatcgccgggactccggagctatgtgcctctt	621
Qy	541	acncayacnattgagrtwacngntaycaymgngngwnsmwmswtatgtnccnytn	600
Db	622	gtctattccagctcagccttcacattactgaccaggtgctttctccgtgagcgtgcc	681
Qy	601	gcnaywsnswnsngnttacyacnathcngaycargtnccnttymwsngtnwsngtnsw	660
Db	682	cagtgccggccttggatggaggaaacagcacttctcagaatacagcctctgaacctt	741
Qy	661	carytnmgcnctngaycgcnggnaayaarcaytttymtmgnaaycarnccnytnacntt	720
Db	742	ggcctccagctccatgaccccagtggtctatctggctgaagctgagcctctcctaacctg	801
Qy	721	gcnctncarctyncaycgcnwsngntayytngcngargcngayytwnsntayacntgg	780
Db	802	gactttggagacagtagtgaaacctgatctctcgggcaecttggtcactcatactac	861
Qy	781	cattyggcnaywsnswngcnacnytnathsmwngcnctyngtngtnwacncaycantay	840
Db	862	ctgagcctggccagctcaetgcccagtggtctcgtcaggctgccattctcctcaect	921
Qy	841	ytngarccngcngcngtnwacngcnargtngtynrcargcnngcnathccnctynacnwsn	900
Db	922	tgtgctctctccccagttccaggcacccagatgggcacaggccaaactgcagaggccct	981
Qy	901	tgycgnwsnswncngtncnccngnacnagcaycgncaymcncnagcngcngargcnccn	960
Db	982	aacaccagctggccagtgccctactacagaagttgtgggtactacacctggtcaggcg	1041
Qy	961	aayacnacngcngcngargtnccnacnacnargtngtnngnacnacnccngcngcargcn	1020
Db	1042	ccaactgcagagccctctggaaccaatctgtgcaggtgcccaaccactgaagtcatagc	1101
Qy	1021	ccnacngcnargcnwsgnagcnacnawsnctncargtnccnacnacnargcngtnathwsn	1080
Db	1102	actgcaactgtgcagatgccaaactgcagagagcacaggtatgacacctgagaagtgcca	1161
Qy	1081	acmgcnctncaratccnvaocngcnarhswnacngnatgacncnngararagtnccn	1140
Db	1162	gtttcagaggtcatgggtaccacactggcagatgtccaactccagaggtcacagttatg	1221
Qy	1141	gtwnsngartnatggcnacnacntnctncngarbtgwnacnccngcngcnacngnatg	1200
Db	1222	acaactgcagaggtatcaatttgtgtcttctggaaccaactgcacaggtaccaact	1281
Qy	1201	acnccngcnargctnwsnathgtngtntymwsngnacnacnccngcngcargtnacnccn	1260
Db	1282	acagagtggtggagaccacagctagagagctacctatccctgagcctgaaggtccagat	1341
Qy	1261	acngarttgggtgnaracnccngcnargtynccnathccngarccngargcngcnag	1320





## source

```
1..2026
/clone="hfl12-2"
/clone_lib="cdm8"
/organism="Homo sapiens"
/cell_line="melanoma H3606"
7..78
7..1995
/standard_name="melanoma-associated ME20 antigen"
/notes="NCBI gi: 494940"
/codon_start=1
/product="me20m"
/db_xref="PID:g494940"
/translation="MDLVKRLHLHLAVIGALLAVGATKVPKPNQWLGVSRQLRTKAW
NRQLYPEATEARLDKRWQGVSLKVSNDGPFLI CANASFSLALNFGSKVLPDQGV
IWNNTLINGSVMGQVPYQETDDACIFDPGCPGSPGSOKRSFVYVVKMTQGVW
QVLGGVSLGIGTGAMLGTHMEVTVYHRRGSRVSLAHSSATITIDQVPFSVS
VSQRLALDGNKHF LRNQPLTFALQLHDPSSYLAEADLSYTWDFGDSSTGLSRALVY
THYLEPGVTAQVVLQAAIPLTSCGSPVPCTTDRHPTAEAPNTTAAQVPTTEWVG
TTPQAPTAEPSTGTVQVPTTEVISTAPVQMPATAESTGMTPEKVPVSEVMGTTIAEM
STPEATGMTPAEVSIVLSGTTAAQVTTTEWVETARELP IPEPGPDASSIMSTESI
TGSIGPLDGTALRLKRVQPLDCLVLYRGSFVTLIDIVOGIESAEILQAVPSEGD
AFELYVSCGGKPEAKMEISSPGCPQAPRLCQVLPSPACQVLVHLIKGSGCTYC
LNVSLADTNSLAVYSTQLIMPQEGAGLGQVPLIVGLVILVMAVYLASLIYRRRLMKQ
DFSVPQLPSSSHWLRLPRIFCSP IGENSPILLSQQV"
79..1992
mat_peptide
/product="me20m"
BASE COUNT 437 a 564 c 564 g 461 t
ORIGIN
```

```
Query Match 95.4%; Score 7274; DB 52; Length 2026;
Best Local Similarity 56.9%; Pred. No. 0.00e+00;
Matches 1129; Conservative 416; Mismatches 437; Indels 3; Gaps 1;

Db 7 atggatctgtgctaaaaagatgccttcttcatttgcgtgtgatagggtgcttgcgtgcct 66
|||||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 1 ATGGAYTNGTNTYNAARMGNTGYTNTYNCAYTNGCNGTNAHGGNGYNTYTNCGN 60

Db 67 gtgggggctacaaagataccgaagaaccaggactggctgtgctcagaggcaactcaga 126
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 61 GTNGGNGCNAARAGTNCNMGNAAYCARGAYTGGYTNNGTNGTNSMNGNCARYTNMGN 120

Db 127 accaaagctggaacaggcagctgtatccagagtggaacagaagccagagacttgactgc 186
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 121 ACNAAACGNTGGNAYMGNCARYTNTAYCNGARTGGACNGARCCNCARMGNTNGAYTGY 180

Db 187 tggagagtggtcaagtgctccctcaagtcagtcagtcagtcagtcagtcagtcagtcagtc 246
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 181 TGGMNGGNGNCARGTNGTNSNTYNAAGTNSNAAYGAYGNCNCAACNTYNAHGGNGCN 240

Db 247 atgctctctctctattgcttgcctccctggaagccaaaggtattgctcagatggg 306
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 241 AAYGCNNSNTTYNSNATHGNTYNTYTCNNGNNSNCARAAARGTNTYTCNCGNAYGN 300

Db 307 caggttatctgggtcaacaataaccatcataatggagagcagggtgtgggagagacagcca 366
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 301 CARGTNAHTGGTNAAYAAACNATHAAAYGNGSNCAARGTNTGGGGNGNCARCN 360

Db 367 gtgtatccccaggaactgacgatcctccttccttccttccttccttccttccttccttcct 426
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 361 GNTAYCCNCAARGAARCAAGAYGAYCNGTGTATHTTYCCNAYGCGNGGNCNTGYCCNWSN 420

Db 427 ggtcttgggtctcagaagaagacttttatgtctggaagacctggggccaactagg 486
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 421 GGNWSNTGGWSNCARAAARMGNWSNTTYGTNTAYGTNTGGAARACNTGGGGNCARTAYTGG 480
```

```
Db 487 caagttcagggggcccaagtgtctggctgagcattgggacaggcaggccaatgctgggc 546
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 481 CARTYTYTNGGNGGNCNGTNGSNCGYTNWSNATHGNCAGNGMNGCNAATGYTNGN 540

Db 547 acacacacatggaagfgactgtctaccatcgccgggggatacccgagcgtatgtgcctct 606
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 541 ACNAYACNATGGARTNACNGTNTAYAYMGNMGNNGNWSNNTAYTGTNCCNTYN 600

Db 607 gctcatccagctcagccttcacacattactgaccaggtgctttctccgtgagcgtgtcc 666
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 601 GCNCAYAWSNWSNCGNTTYACNATHACNGAYCARGTNCNTTYWSNGTNSNGTNSN 660

Db 667 cagttgcgggcttggatgaggggacaacacttcctcgaaatcagcctctgcactct 726
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 661 CARYTNMGNCYNTGAYGCGNGNAAYAAACAYTTYTNMGNAAYCARCCNYTNACNTY 720

Db 727 gctctccagctccatgaccccaagtgcctatctggctgaagctgacctctcctacacctgg 786
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 721 GCNYTNCARTYNCAYTAYCWNWSNGNTAYTNGCNGARGCNGAYTNWSNTAYACNTGG 780

Db 787 gactttgagacagtagtgaacctgatctctcggcacttgtgtcactcatacttac 846
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 781 GAYTTYGNGAYWSNWSNCGNTYATNATHWSMNGNGYTNATNCARGCNGNATHCCNYTNACN 840

Db 847 ctggagctggcccaagtcaactgcccaggtgtcctgcaaggtgcccattcctcaccctcc 906
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 841 YTNARGCNGCNGCNGTNAACNGNCARCTNGTNTNCARGCNGCNGNATHCCNYTNACNWSN 900

Db 907 tgtggtctctcccagttccaggaccacagatgggcacaggcccaactgcagaggccct 966
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 901 TGYGGSNWSNCCNGTNCNGGNACNACNGAYGNCAYMGNCCNACNGCNGARGCNGCN 960

Db 967 aacacacagctggcgaagtgcctactacagaagttgtgggtactacacctggtcaggcg 1026
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 961 AAYACNACNGCNGCARGTNCNACNACNGARGTNGTNGNACNACNCCNGCNGCARGCN 1020

Db 1027 ccaactcgagagccctctggaaccaactctgtgcaagtgccaaccactgaagtcataagc 1086
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 1021 CCNACNGCNGCNGCNGNWSNCGNACNWSNCTNCARGTNCNACNACNACNGARGTNAHWSN 1080

Db 1087 actgcaactgtgcagatgccaactgcagagacagaggtatgacacctgagaagtgcca 1146
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 1081 ACNGCNGCNGTNCARATGCCNACNGCNGARMSNACNCGNATGACNCCNGARAARGTNCN 1140

Db 1147 gtttcagaggtcatgggtaccacactggcagagatgtcaactccagaggtcacaggtatg 1206
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 1141 GTNWSNARGTNTATGGGNACNACNTYTCNGARATGMSNACNCCNCGCNGCNAACNGNATG 1200

Db 1207 acactcgagaggtatcaattgtggtgttcttctggaaccacagctgcacaggttaacaact 1266
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 1201 ACNCCNGCNGARGTNSNATHGNTNGTNTYTNWSNNGNACNACNCCNGCNGCARGTNAACN 1260

Db 1267 acagagtggttgagaccacagctagagagctacctatccctgagcctgaaggtccagat 1326
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 1261 ACNGARTGGGTNGARACNACNCGNCGNARGYTNCNATHCCNARGCNGCNGCNGAY 1320

Db 1327 gccagctcaatcatgtctacggaaagattacaggttccctgggcccctgctggtggt 1386
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 1321 GCNWSNNSNATHATGWSNACNGARMSNATHACNGNWSNYTNGCNGCNYTNTYTNAYGNGN 1380

Db 1387 acagccaccttaagctgggtgaagagacagctccctcctggtggtgtctgtatcgatat 1446
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 1381 ACNCGNACNTYTNMGNTNGTNAARMGNARGTNCNCTYTNAYTGYGTNTYTNAYMGNATY 1440
```













Db	490	ggggggtccaggtccaggtcagcatctcactacgcacgaagctgggcacacaca	549
Qy	490	GGGGGCGCTGWSNGNYTNWSNATHGNCAGNMGNGCNGATGYTNGNACNACYN	549
Db	550	atggaagtgaactctaccacgcaggggttccacagatcagctgccccttgctcagcc	609
Qy	550	ATGGAGTGNACNGTNTATCAATMGNGMGNGWSNMTATGTNCNCTYNGCNACYN	609
Db	610	agttcaacctcaactaccagcaggtacccttctccgtgagtggtccagctcacag	669
Qy	610	WSNWSNGNTTYACNATHACNGAYCARGTNCNTTYWSNGTNSNGTNMSNCARYTN	669
Db	670	gccttgagcggagagacacacagcactctcctgagaatcatcctcactcttgccctcag	729
Qy	670	GCNYTNGAYGGNGNAAATCAATYTYTNMGNAATYTCARCCNCTYTNACNTTYG	729
Db	730	ctccagcaccacagtggttatgttggcggagcgcagcactctcctacacatggagcttgg	789
Qy	730	YTNCAATGAYCWNWSNGNTATYTNCGNARGCNGAYTNMSNTATACNTGGATYTN	789
Db	790	gatggtactggagaccctgatctctcgggcacttgatgcactcacacttaacctggagctg	849
Qy	790	GATWSNWSNGNACNYTNATHWSNMGNCNTYTNCTGNTGNTGNTGNTGNTGNTGNTG	849
Db	850	ggctcagtcactgccacaggtggtactgcaggtggcattcctctgttctcctgtgtcc	909
Qy	850	GGNCCGNTGACNGCARGTNGTNTGNTGNTGNTGNTGNTGNTGNTGNTGNTGNTGNTG	909
Db	910	tcccaggtcccgggtaccacagatggtcactgcacactgcagaagcactggcaacaca	969
Qy	910	WSNCCGNTGNCNGNACNACNGAYGNCAYMGNCNACNGCNGCNGCNGCNGCNGCNGC	969
Db	970	tctaggcagaagcaacctacacaaagttdtgggtactacactggccagatgccacataca	1029
Qy	970	GCNGNCARGTNCNACNACNGARGTNGTNGNACNACNCGNCGNCGNCGNCGNCGNCGN	1029
Db	1030	cagccctctgggaaccacagttgtcaaaatgccaacacagaggtcacagctactacatct	1089
Qy	1030	GARCCWSNCGNACNACNWSNGTNCARTGNTCCNACNACNGARGTATWSNACNCGN	1089
Db	1090	gaqcagatgctgacctcagcggctcatagatcaccaactggcagaggtgtcaactacagag	1149
Qy	1090	GTNCARATGCCNACNCGNARGWSNACNGNATGACNCGNARGAARGTNCNCTGNTWSN	1149
Db	1150	ggtcaggtaccacac	1165
Qy	1150	GTNATGGNACNACNY	1165
RESULT	8		
LOCUS	BOVBRPE1A	1584 bp	MAM
DEFINITION	Bovine retinal pigment (RPE1) mRNA,	3' end.	08-MAR-1993
ACCESSION	M81193		
NID	q162770		
KEYWORDS	Retinal pigment.		
SOURCE	Bos taurus (library: Unizap XR) calf retinal pigment epithelium cDNA to mRNA.		
ORGANISM	Bos taurus		
REFERENCE	1 (bases 1 to 1584)		
AUTHORS	Kim, R.Y. and Wistow, G.J.		
TITLE	The cDNA RPE1 and monoclonal antibody HMB-50 define gene products		
FEATURES	1 (bases 1 to 1584)		
FEATURES	1 (bases 1 to 1584)		

[illegible]

D	b	482	gaggtcctggaaccacagctggccaaagtgcctactacaagatcatgggccacacct	541
Q	y	952	GARGCNCNAYACNAACNGGNCARGTTCNNACNACNGARGTNTGNACNACNCCN	1011
D	b	542	ggccaagtgcacaactgcagaggcccctggcaccacagtgggtggcgcaaccacagag	601
Q	y	1012	GCNCARGCNCAACNGCNGARGCNCNSNGNACNACNWSNCTNCARTGCCNACNACNGAR	1071
D	b	602	g--at-gta-g---gtacca-----cac-ctgaq-caggtggca--acctcc---	637
Q	y	1072	GTNATHWSNACNCCNCGTNCARTATGCCNACNCGNGARHWSNACNGENATGAONCNGAR	1131
D	b	638	aaaft--c--t-t---aaqt-ac----cac-ac-c-agtggagatgccaactgcaaacgt	682
Q	y	1132	AARGTNCCGTNWSNARGRTNATCGGNACNACTYTNCCNGGARATGWSNACNCCNCGANG	1191
D	b	683	acagtgaaagacactg---aagtgtcaactacagagccctctggaaccacagttacaag-	739
Q	y	1192	ACNGGNATCAACNCCNGARGTNWSNATHGTGTNYTNWSNGNACNACNCGNCGNCAR	1251
D	b	740	ggaacaactccagagctggtggagaccacagctggagaggtgtccactcctgagctgcg	799
Q	y	1252	GTNACNACNARGRTGGGTNGARAACNACNCGNMGARITYTNCNATHCCNGARCNGAR	1311
D	b	800	ggttcaaatactagctattcctacagaaggtgaetcgaggtccccctgagtcctctg	859
Q	y	1312	GMCCNGAYGCNWSNWSNATHGWSNACNARGHSNATHCNGSNWSTYTNCCNCTYN	1371
D	b	860	cggatgacactgcacacttagtctggagagcgccagccccctggattgttgtctg	919
Q	y	1372	YTNGAYGNCNCCNACTYTMNGTYTNGTAARMGRCARGTNCCTYNGAVTGYGTNYTN	1431
D	b	920	tatcgctatggctcttttccctcacctggacattgt-cag--tatgagagtgctgag	976
Q	y	1432	TAYMGNATYGONSNTYWSNGTNAACNVTNGAVATHGTCNARGNATHGARNWSNCGAR	1491
D	b	977	atcctacaggtctgtcatccagtgaaggagatgcatcttgagctgactgtgctctgcac	1036
Q	y	1492	ATHYTNARGCNTNCNWSNCGNGARGGAGYGCNTTYGARYTNACNCTNWSNTGYCAR	1551
D	b	1037	ggcgggtaccagaagacctgcattgacatctcatcgccaggggtgcagctgcctgcc	1096
Q	y	1552	GGNGNYTNCNAARCAAGCNTGYATGGARATHWSNCCNGGNTGYCARCCNCCNGN	1611
D	b	1097	cagcggtgtgcagctgtgcccccagccagcagctgccagctggtttgcaccaggtta	1156
Q	y	1612	CARMGNITGYCARCNGTYTNCNWSNCCNGONTGYCARYTNGTNTYNCACARATH	1671
D	b	1157	ctgaaggtggctcaggaagacctactgcctcaatgtctttggctgatgccaatagctg	1216
Q	y	1672	YTNAARGNGNWSNCAACNTATGYTINAAYGTNWSNYTNGCNAYACNAAVWSNYTN	1731
D	b	1217	ggcatggtcagcaccacctgtcatgctggccaagaagcagacctcagcgagctcct	1276
Q	y	1732	CGNGTNGTNSNACNARYTHATGCCNGNARGRGONGNYTNGNCAARGTNCN	1791
D	b	1277	ctgtctgtgggcatcttgctggctacaactttgtgcttgcactctgatatacagg	1336
Q	y	1792	YTNAHTGTNGGNATHYTNNGTYTNGTYTGCGTNGTNGTYTNGCNWSNTYTHATYMG	1851
D	b	1337	cgaagaactatgaagcaagacctcagaagtccecttccccagctgcccacaggtagaacc	1396
Q	y	1852	MGNNGTYNATGAARCATGYT-----TYMSNTGTCNCCNARTYTNCCNCAVWSNWSN	1905

D	b	1397	cagtggctacgtccctcggtgcttcgcgcctcttgcgccattggtgagacaaccctc	1455
			:   :        :     :     :     :     :     :     :	
Q	y	1906	CATGGTGMGNVTNCNMGNATHTTTGTGMSNTGYCCNATHGCNGARAAWNSNCNVTN	1965
D	b	1457	ctcagtggaacgagcgtt 1473	
			:   :        :     :	
Q	y	1966	YTWWSNGNCARCART 1982	
			:	
<hr/>				
		RESULT	9	
ID		HSMEL103	standard; DNA; PRI; 2001 BP.	
AC		U31797;		
DT		23-DEC-1995 (Rel. 46, Created)		
DT		23-DEC-1995 (Rel. 46, Last updated, Version 1)		
DE		Human melanocyte protein Pmel 17 gene, exons 4, 5, and 6.		
KW		.		
OS		Homo sapiens (human)		
OC		Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;		
OC		Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.		
RN		[1]		
RP		1-2001		
RA		Bailin T., Lee S.T., Spritz R.A.;		
RT		"Genomic organization and sequence of D12S53E (Pmel 17), the human		
RT		homologue of the mouse silver (sl) locus";		
RL		Unpublished.		
RL		[2]		
RP		1-2001		
RA		Bailin T.;		
RT		?		
RL		Submitted (17-JUL-1995) to the EMBL/GenBank/DBDJ databases.		
RL		Richard A. Spritz, Medical Genetics, 445 Henry Mall, University of		
RL		Wisconsin, Madison, WI 53706, USA		
CC		NCBI gi: 1125059		
FH		Key	Location/Qualifiers	
FT		source	1..2001	
FT		/organism="Homo sapiens"		
FT		/chromosome="12"		
FT		/map="12q13-q14"		
FT		/note="DSEG number: D12S53E"		
FT		exon	95..229	
FT		/number=4		
FT		exon	623..781	
FT		/number=5		
FT		exon	1030..1755	
FT		/number=6		
SQ		Sequence 2001 BP; 457 A; 554 C; 492 G; 498 T; 0 other;		
<hr/>				
		Query Match	34.9%; Score 2662; DB 10; Length 2001;	
		Best Local Similarity	58.4%; Pred. No. 0.00e+00;	
		Matches	429; Conservative 130; Mismatches 176; Indels 0; Gaps	
D	b	1022	caatccagcacaggctcttcctcgtagcgctgccagttgcgcgcttgatgagag 1081	
			:	
Q	y	623	CNATHCAGTCARGTNCNTTWSNGTWSNGTWSNCARYTNMNGCNTNGAYCGNG 682	
D	b	1082	ggaacaagcattcctcgagaatcagcctctgaccttgcctccagctccatgaccca 1141	
Q	y	683	GNAAYAACAATTYYTNMGNAAYCARHCNTYNACNTTYGCNYTCNARYNCAYCCNW 742	
D	b	1142	gtggctatctggctgaagactgacctctctacacctggaccttggagacagtgagaa 1201	
			:     :     :	
Q	y	743	SNGETATYINGNCARGCAGGAYTNWNTAYACNTGGGAYTTGGNGAYWSNMGNA 802	







```
Qy 1283 CMMNGARYTCCNATHCCGARGCNGCNGCNGAYCWNWSNATHATGKSNACNG 1342
      | : | | : | | | : | | | : | | | : | | | : | | | : | | | : | | |
Db 1742 aaagtattacagta 1756
      | : : | | : | | :
Qy 1343 ARWSNATHACNGNW 1357

RESULT 11
ID HSMEL04 standard; DNA; PRI; 1559 BP.
AC U31798;
DT 23-DEC-1995 (Rel. 4.6, Created)
DT 23-DEC-1995 (Rel. 4.6, Last updated, Version 1)
DE Human melanocyte protein Pmel 17 gene, exons 7, 8, 9, and 10.
KW .
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homiidae.
RN [1]
RP 1-1559
RA Bailin T., Lee S.T., Spritz R.A.;
RT "Genomic organization and sequence of D12S53E (Pmel 17), the human
RT homologue of the mouse silver (si) locus";
RL Unpublished.
RN [2]
RP 1-1559
RA Bailin T.;
RT ;
RL Submitted (17-JUL-1995) to the EMBL/GenBank/DBJ databases.
RL Richard A. Spritz, Medical Genetics, 445 Henry Mall, University of
RL Wisconsin, Madison, WI 53706, USA
CC NCBI gi: 1125060
FH Key Location/Qualifiers
FT source 1..1559
FT /organism="Homo sapiens"
FT /chromosome="12"
FT /map="12q13-q14"
FT /note="DSEG number: D12S53E"
FT exon 382..496
FT /number=7
FT exon 736..821
FT /number=8
FT exon 932..1135
FT /number=9
FT exon 1238..1327
FT /number=10
SQ Sequence 1559 BP; 340 A; 392 C; 413 G; 414 T; 0 other;

Query Match 10.2%; Score 780; DB 10; Length 1559;
Best Local Similarity 55.8%; Pred. No. 1.74e-90;
Matches 116; Conservative 48; Mismatches 44; Indels 0; Gaps 0;

Db 930 ggctgcccaaggaagcctgcagtgagatctcatccaggggtgccagccccctgccagc 989
      | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Qy 1556 GNYTNCNNAARGCNGTGYATGGARATHWSNWSNCCNGTGYCARCNCNGCNCARM 1615

Db 990 ggctgtgccagcctgtgtaccagccagcctgccagctggcttgcaccagatactga 1049
      | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Qy 1616 GNYTNGYCARCNGTNYTNCNWSNCCNGTGYCARYTNGTNYTNCAYCARATHYNA 1675

Db 1050 aggggtggctggggacatactgctcaatgtctctgtggtgataccaacagcctggcag 1109
      | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Qy 1676 ARGNGGNWSNGNACNTATTGYTTNAAAYGTNWSNYTNGCNGAYACNAAYSNYTNGCNG 1735
```

```
Db 1110 tggtcagcaccagcttatcatgcctgg 1137
      | : | : | | : | | : | | : | | : | | : | | : | | : | |
Qy 1736 TNGTWNWACNARYTNATHATGCCNGG 1763

RESULT 12
LOCUS HSPMEL04 1559 bp DNA PRI 19-DEC-1995
DEFINITION Human melanocyte protein Pmel 17 gene, exons 7, 8, 9, and 10.
ACCESSION U31798
NID g1125060
KEYWORDS .
SEGMENT 4 of 5
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1559)
AUTHORS Bailin,T., Lee,S.-T. and Spritz,R.A.
TITLE Genomic organization and sequence of D12S53E (Pmel 17), the human
TITLE homologue of the mouse silver (si) locus
JOURNAL Unpublished (1995)
REFERENCE 2 (bases 1 to 1559)
AUTHORS Bailin,T.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-1995) Richard A. Spritz, Medical Genetics, 445
Henry Mall, University of Wisconsin, Madison, WI 53706, USA
COMMENT NCBI gi: 1125060
FEATURES
source Location/Qualifiers
      1..1559
      /organism="Homo sapiens"
      /chromosome="12"
      /map="12q13-q14"
      /note="DSEG number: D12S53E"
      exon 382..496
      /number=7
      exon 736..821
      /number=8
      exon 932..1135
      /number=9
      exon 1238..1327
      /number=10
BASE COUNT 340 a 392 c 413 g 414 t
ORIGIN

Query Match 10.2%; Score 780; DB 86; Length 1559;
Best Local Similarity 55.8%; Pred. No. 1.74e-90;
Matches 116; Conservative 48; Mismatches 44; Indels 0; Gaps 0;

Db 930 ggctgcccaaggaagcctgcagtgagatctcatccaggggtgccagccccctgccagc 989
      | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Qy 1556 GNYTNCNNAARGCNGTGYATGGARATHWSNWSNCCNGTGYCARCNCNGCNCARM 1615

Db 990 ggctgtgccagcctgtgtaccagccagcctgccagctggcttgcaccagatactga 1049
      | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Qy 1616 GNYTNGYCARCNGTNYTNCNWSNCCNGTGYCARYTNGTNYTNCAYCARATHYNA 1675

Db 1050 aggggtggctggggacatactgctcaatgtctctgtggtgataccaacagcctggcag 1109
      | : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Qy 1676 ARGNGGNWSNGNACNTATTGYTTNAAAYGTNWSNYTNGCNGAYACNAAYSNYTNGCNG 1735

Db 1110 tggtcagcaccagcttatcatgcctgg 1137
      | : | : | | : | | : | | : | | : | | : | | : | | : | |
Qy 1736 TNGTWNWACNARYTNATHATGCCNGG 1763
```

```

RESULT 13
LOCUS HSPMEL02 673 bp DNA PRI 19-DEC-1995
DEFINITION Human melanocyte protein Pmel 17 gene, exons 2 and 3.
ACCESSION U31807
NID g1125058
KEYWORDS
SEGMENT 2 of 5
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 673)
TITLE Genomic organization and sequence of D12S53E (Pmel 17), the human
homologue of the mouse silver (sl) locus
JOURNAL Unpublished (1995)
REFERENCE 2 (bases 1 to 673)
AUTHORS Spritz, R.A.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-1995) Richard A. Spritz, Medical Genetics, 445
Henry Mall, University of Wisconsin, Madison, WI 53706, USA
COMMENT NCBI gi: 1125058
FEATURES
source
1..673
/organism="Homo sapiens"
/chromosome="12"
/map="12q13-q14"
/notes="DSEG number: D12S53E"
128..241
/number=2
400..546
/number=3
186 a 142 c 179 g 166 t
exon
exon
BASE COUNT 186 a 142 c 179 g 166 t
ORIGIN
Query Match 7.5%; Score 570; DB 86; Length 673;
Best Local Similarity 56.0%; Pred. No. 9.74e-58;
Matches 89; Conservative 37; Mismatches 33; Indels 0; Gaps 0;
Db 387 ttcttccagggtggtcaaggtccctcaaggtcagtgatgaggtcctacactgattgg 446
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 177 YTGCTGCGNGGNGCARGCTNWSNYTNAARGTNWSNAAAYCGNCCNACNYTNATHGG 236
Db 447 tgaatgcctctctctattgccttgaactccctggaagcgaaggaaggtattgccaga 506
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 237 NCGNAAACGNWSNTTWSNATHGCONYTNAAATTTCCGNGNWSNCARAAAGTNTNCCNGA 296
Db 507 tggcagggttatctgggtcaacaataccatcatcaatgg 545
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 297 YGCNCARGTNATHTGGGTNAAYAAACNATHATHAAYGG 335

RESULT 14
ID HSPMEL02 standard; DNA; PRI; 673 BP.
AC U31807;
DT 23-DEC-1995 (Rel. 46, Created)
DT 23-DEC-1995 (Rel. 46, Last updated, Version 1)
DE Human melanocyte protein Pmel 17 gene, exons 2 and 3.
KW
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominidae.

```

```

RN 11
RP 1-673
RA Bailin T., Lee S.-T., Spritz R.A.;
RT "Genomic organization and sequence of D12S53E (Pmel 17), the human
homologue of the mouse silver (sl) locus";
RL Unpublished.
RN 12
RP 1-673
RA Spritz R.A.;
RT
RL Submitted (17-JUL-1995) to the EMBL/GenBank/DBJ databases.
RL Richard A. Spritz, Medical Genetics, 445 Henry Mall, University of
RL Wisconsin, Madison, WI 53706, USA
CC NCBI gi: 1125058
FH Key Location/Qualifiers
FT source
1..673
/organism="Homo sapiens"
/chromosome="12"
/map="12q13-q14"
/notes="DSEG number: D12S53E"
128..241
/number=2
400..546
/number=3
Sequence 673 BP; 186 A; 142 C; 179 G; 166 T; 0 other;
Query Match 7.5%; Score 570; DB 10; Length 673;
Best Local Similarity 56.0%; Pred. No. 9.74e-58;
Matches 89; Conservative 37; Mismatches 33; Indels 0; Gaps 0;
Db 387 ttcttccagggtggtcaaggtccctcaaggtcagtgatgaggtcctacactgattgg 446
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 177 YTGCTGCGNGGNGCARGCTNWSNYTNAARGTNWSNAAAYCGNCCNACNYTNATHGG 236
Db 447 tgaatgcctctctctattgccttgaactccctggaagcgaaggaaggtattgccaga 506
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 237 NCGNAAACGNWSNTTWSNATHGCONYTNAAATTTCCGNGNWSNCARAAAGTNTNCCNGA 296
Db 507 tggcagggttatctgggtcaacaataccatcatcaatgg 545
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 297 YGCNCARGTNATHTGGGTNAAYAAACNATHATHAAYGG 335

RESULT 15
LOCUS HSPMEL05 1178 bp DNA PRI 19-DEC-1995
DEFINITION Human melanocyte protein Pmel 17 gene, exon 11 and partial cds.
ACCESSION U31799
NID g1125061
KEYWORDS
SEGMENT 5 of 5
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1178)
AUTHORS Bailin, T., Lee, S.-T. and Spritz, R.A.
TITLE Genomic organization and sequence of D12S53E (Pmel 17), the human
homologue of the mouse silver (sl) locus
JOURNAL Unpublished (1995)
REFERENCE 2 (bases 1 to 1178)
AUTHORS Bailin, T.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-1995) Richard A. Spritz, Medical Genetics, 445

```

Henry Mall, University of Wisconsin, Madison, WI 53706, USA

COMMENT	NCBI gi: 1125061

## FEATURES

```
source
1..1178
```

/organism="Homo sapiens"

/chromosome="12"

/map="12a13-a14"

/note="DSEG number: D12S53E"

join(U31808:158..232,U31807:128..241,U31807:400..546,

U31797:95..229,U31797:623..781,U31797:1030..1755.

U31798:382..496,U31798:736..821,U31798:932..1135,

U31798:1238..1327,298..432)

/note="qp100; ME20; NCBI qi: 1125063"

```

/codon start=1

```

/product="melanocyte protein Pmel 17"

```
/db xref="PID:q1125063"
```

\_&translation="MDIVLKERCLLHLAVIGALLAVCANATKPRQNDPLCVSRQLRTKAN  
 NQDNLTPYETEDAQRLDQWRQGSYLVNSDQFGLTIGANFSAITANPFGSKQVLPDQGV  
 YWNTWNTIINGSGWCGQVQYPTQDADICFGDGCPSGMSOKRSFVYVWKVQGYW  
 QVQLVGPVSGSLGTCRAMLGTHMEVTVYHRSGSYVLAHSSAFTITDQVPSVS  
 SYLRAIDGNGKHLRNQPLITLALQHDQSGYLAEADLYWFOGDUSGTLISRALV  
 YHTIYLPFGVTAQVLAQALITLSCGSSPVQDTCRHPATATAGQVLPVTEVG  
 QVQDQVQDQVQVTEVLTSTAPVQMPATAESTGPEKPVSEVWGCTITAEI  
 STTPKATQMPAEVSLVSLTGAQVQVLTQEWTELTARELPIDPEQGPASSIMTESI  
 TCGESGLPLDGTATILVKRQVQDLYCLVYRGSFVTLIDVQGETAEILQAVPSGEGD  
 AFELTVISCGGGLPEAKQWELSPGCGQVLPQVLPSPACQVLIHQLIKGSGTYC  
 NVNLSIATNSLAVSLQVLMIPQGEALQGVPLVQLGVLVMAVVLASLITYRRRLMKQD  
 FSVQLPEHSSHSILRLRPIFCSPQTCGNSPLLSGQGV"

BASE COUNT	298 a	231 c	373 a	276 t
------------	-------	-------	-------	-------

## ORIGIN

Query Match 6.7%; Score 511; DB 86; Length 1178;

Best Local Similarity 46.6%; Pred. No. 8.65e-49;

Sequence	Matches	Conservative	Mismatches	Indels	Gaps
1	69	69	46	33	0
2	69	69	46	33	0
3	69	69	46	33	0
4	69	69	46	33	0
5	69	69	46	33	0
6	69	69	46	33	0
7	69	69	46	33	0
8	69	69	46	33	0
9	69	69	46	33	0
10	69	69	46	33	0
11	69	69	46	33	0
12	69	69	46	33	0
13	69	69	46	33	0
14	69	69	46	33	0
15	69	69	46	33	0
16	69	69	46	33	0
17	69	69	46	33	0
18	69	69	46	33	0
19	69	69	46	33	0
20	69	69	46	33	0
21	69	69	46	33	0
22	69	69	46	33	0
23	69	69	46	33	0
24	69	69	46	33	0
25	69	69	46	33	0
26	69	69	46	33	0
27	69	69	46	33	0
28	69	69	46	33	0
29	69	69	46	33	0
30	69	69	46	33	0
31	69	69	46	33	0
32	69	69	46	33	0
33	69	69	46	33	0
34	69	69	46	33	0
35	69	69	46	33	0
36	69	69	46	33	0
37	69	69	46	33	0
38	69	69	46	33	0
39	69	69	46	33	0
40	69	69	46	33	0
41	69	69	46	33	0
42	69	69	46	33	0
43	69	69	46	33	0
44	69	69	46	33	0
45	69	69	46	33	0
46	69	69	46	33	0
47	69	69	46	33	0
48	69	69	46	33	0
49	69	69	46	33	0
50	69	69	46	33	0
51	69	69	46	33	0
52	69	69	46	33	0
53	69	69	46	33	0
54	69	69	46	33	0
55	69	69	46	33	0
56	69	69	46	33	0
57	69	69	46	33	0
58	69	69	46	33	0
59	69	69	46	33	0
60	69	69	46	33	0
61	69	69	46	33	0
62	69	69	46	33	0
63	69	69	46	33	0
64	69	69	46	33	0
65	69	69	46	33	0
66	69	69	46	33	0
67	69	69	46	33	0
68	69	69	46	33	0
69	69	69	46	33	0
70	69	69	46	33	0
71	69	69	46	33	0
72	69	69	46	33	0
73	69	69	46	33	0
74	69	69	46	33	0
75	69	69	46	33	

Db 281 cttctcccaatatcagqccagacttatgaagcaagacttctccgtataccccagttgccac 340

[illegible]

1835 CNWSNYTNAHTAYMGNMGNMNYTNAATGAARCARGAYTTYWSNGTNCNCARYTNCNC 1894

Db 341 atagcagcagtcactggctgcgtctacccgcacatcttctactcttctccattgattgaga 400

1895 AYWSNWSNCAYTGGYTNMGNYTNCNMGNAHTTYTGYWSNTGYCCNATHGGNGARA 1954

Db 401 atagccctcctcagtqqcacaggt 428

	::	
	::	
	::	
	::	

1955 AYWSNCCNYTNYTWSNGNCARCARGT 1982

Search completed: Tue Aug 27 19:08:20 1996

Job time : 1321 secs.

(MT)

Result	No.	Score	Query			ID	Description	Pred. No.
			Match	Length	DB			
	1	7374	96.7	2172	17	T027116	MART-1 melanoma anti	0.00e+00
	2	7364	96.6	2115	16	Q96055	Sequence encoding mel	0.00e+00
	3	6699	87.9	2131	17	T03760	Melanoma-specific imm	0.00e+00
	4	177	2.3	1047	2	Q10572	Human Natriuretic Pep	3.52e-02
c	5	173	2.3	2351	2	Q10498	B-amyloidquefaciens s	6.56e-02
c	6	157	2.1	1065	17	Q94335	Degenerate Alteromora	7.46e-01
	7	150	2.0	36	16	T05466	Sequence encoding imm	2.10e+00
	8	147	1.9	755	1	Q03970	Clone P25b	3.24e+00

DR P-PSDB; R84854.

PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.

PS Disclosure; Fig 4A-4B; 184pp; English.

CC The nucleic acid encodes cDNA25, a melanoma antigen (MART-1)  
 CC which is recognized by T-lymphocytes. cDNA25 is a derivative of  
 CC the melanocyte-melanoma-specific antigen gp100 (see R84855).  
 CC Antigen cDNA25 is a source of immunogenic peptides (see R84199)  
 CC which are optionally modified (see R84200-R84211) and used in  
 CC medicaments, especially vaccines, for the treatment or prevention  
 CC (by immunization) of melanoma. Antibodies against cDNA25 and its  
 CC immunogenic peptides may be used in the detection and isolation  
 CC of the antigen from a sample, the detection of which is indicative  
 CC of a disease state (melanoma or metastatic melanoma).

SQ Sequence 2172 BP; 512 A; 594 C; 578 G; 488 T;

Query Match 96.7%; Score 7374; DB 17; Length 2172;

Best Local Similarity 57.0%; Pred. No. 0.00e+00;

Matches 1130; Conservative 416; Mismatches 436; Indels 0; Gaps 0;

Db 39 atggatctggctaaaaagatgctcttcattgctgtgatgtagtgcttggctgct 98

Qy 1 ATGCAYTNGTNTYNAARMGNTGYTNTYNCAYTTCGCTNATHGCGNCTYNTYTCGN 60

Db 99 gtgggggtcaaaagtaccgaagaaacagagactggtgtgtcctcaaggcaactcaga 158

Qy 61 CTNGGCGNACNAARGTCCNNGNAAYCARGAYTGGYTGNGGTGNTWSMNGNCARYTNMGN 120

Db 159 accaaagcctggaacagcagctgtatccagagtggaagacagccagagacttgactgc 218

Qy 121 ACNAARCGNTGGAAVMGNCARYTNTAYCNGARTGGACNGARCCNCARMGNTYNGAYTGY 180

Db 219 tggagagtggtcaagtgtccctcaagtcagtaatgatggctacactgattggtgca 278

Qy 181 TCGMNGNGNGNCARGTNSWNTNNAARGTNWSNAAYGAYGNCNACNATYNTATHGGNCN 240

Db 279 aatgcctctctctattgcttgaacttccttgaagccaaagattgcccagatggg 338

Qy 241 AAYCGNWSNTTYSNATHGNTYNTAAVTTCGNGNWSNCABAARGTNTYTCNGAYGN 300

Db 339 caggttatctgggtcaacaataccatcatcaatggagccaggtgtgggagacagcca 398

Qy 301 CARGTNATHTGGGTNAAYAAAYACNATHATHAAYGGSWNSNCARGTWTGGGNGNCARCCN 360

Db 399 gtgtatcccccaggaactgacgatgcctgeatcttccctgatggtggaccttgcccactc 458

Qy 361 GTNTAYCCNCARGARACNAYGACGNTGYATHTYTCNGAYGCGNGCNCCTGYCOWSN 420

Db 459 ggccttggtctcagaagaagcttggtttatgtctggaagacctggggcccaactagg 518

Qy 421 GGNWSNTGGSNCARAARMGNSNTTGTNTAYGTNTGGAARACNTGGCNCARTAYTGG 480

Db 519 caattctcagggggccagtgctcgggtgagcattgggacagcagggcaatgctgggc 578

Qy 481 CARTTYTNGGNGCCNGCTNWSNGNTYNTWSNATHGNCACNGMNGNCNATGYTNGSN 540

Db 579 acacacaccatggaagtgactgtacacatcccgccggggtcccgagactatgtgcctctt 638

Qy 541 ACNCAACNATCGARTNACNGTNTAYCAYMGMNGMNGSNMNGNSNTAYGTNCNVTN 600

Db 639 gctcattccagctcagccttcaccattactgaccaggtgctcttctccgtgagcgtgcc 698

Qy 601 GCNCAYSWNSWNSGNCNTTYACNATHACNAYCARGTNCNCTTYWSNGTNSGNTWSN 660

Db 699 cagttgcccgttgatggaggggaacaagactcctcgagaaatcagcctctgacctt 758

Qy 661 CARYTNMGNCNTYNGAYGCGGNAAYAAARCAVTTYTNMGNAAYCARCCNYTNAcntty 720

Db 759 ggcctccagctccatgaccaccagtgctatctggctgaagctgaocctctcctacacctgg 818

Qy 721 GCNTYTCARYTNCAYGAYCCNWSNGNTAYTYTNGCNGARGCNGAYTTHWSNTAYACNTGG 780

Db 819 gactttgagacagtagtggaaacctgatctctcgggcacttgtggtcactcatacttac 878

Qy 781 CAYTTYGNGAYGAYSWNSWNGNACNTYNTATHWSMNGNCTYNTGNTGNTNACNCAACNTAY 840

Db 879 ctggagcctggcccagtcactgcccaagtggtcctcgaggtgcccattcctcctcaccctcc 938

Qy 841 YTNARGCNGCNGCNGTNAACNGCARGTNGTNTYTCARGCNGCNAHCCNYTNAQNWSN 900

Db 939 tgtggtctctcccagttccaggaccacagatgggcacagggcccaactgcagaggccct 998

Qy 901 TGYGNSWNSWNSCNGTNCNGNACNACNGAYGNCAYMGCNCCNACNGCARGCNGCNCN 960

Db 999 aaacacacagctggccaagtgcctactacagaagttgtgggtactacacctggtcaggcg 1058

Qy 961 AAYACNACNGCNGCARGTNCNCAONACNARGTNGTNGNACNACNCCNCGNCARGCN 1020

Db 1059 ccaactgcagagccctctggaaccatctgtgcaggtggcaaccactgaagtcataagc 1118

Qy 1021 CCNACNGCARGCNCWSNGNACNACNWSNGTNCARGTNCNCAONACNACNARGTNAHWSN 1080

Db 1119 actgcaactgtgcagatgccactgcagagacacaggtatgacacctgagaagtgcca 1178

Qy 1081 ACNGCNCNTNCARATGCCNACNGCNGARBSWNAACGNATGACNCCNGARAARGTNCNCN 1140

Db 1179 gtttcagagtggtgggtaccacactggcagagatgtcaactccagaggtcacaggtatg 1238

Qy 1141 GTNWSNARGTNATGGGNCNACNCTYTCNGCNGARATGHSNACNCCNCGARGCNGCNGATG 1200

Db 1239 acactgcagagtgatcaattgtggtcttctggaaccacagctgcacaggtacaact 1298

Qy 1201 ACNCCNGCARGTNSWNTATGCTGNTYTNWSNGNACNACNCGCNGCNGCARGTNAACN 1260

Db 1299 acagagtggtggagaccacagctagagagctacctatccctgagctgaaggtccagat 1358

Qy 1261 ACNGARTGGGTNGARACNACNCGNMGNGARYTNCNATHCCNARGCNGARGCNGCNGAY 1320

Db 1359 gccagctcaatcatgtctacgaaagattacaggttcctctggggccctgctggatggt 1418

Qy 1321 GCNWSNWSNATHGWSNACNARBSNATHACNCGNWSNYTNGCNCNCTYNTYNGAYGN 1380

Db 1419 acagccacttaaggtggtgaagacaaagtcctcctggattgtgtctgtatcgatat 1478

Qy 1381 ACNCGNACNCTYTMNGNTYTNNAARMGNCARGTNCNCTYTCAYTGTNTYNTAYMNTAY 1440

Db 1479 gttcctctttccgtcacctcgacatgtccagggtattgaaagtgcagagatcctcgag 1538

Qy 1441 GGNWSNTTYSNGTNGAYATHGTCNARGGNATHGARMWSNCGNARGARATHYTNCAR 1500

Db 1539 gctgtccgtccogtggaggggagtgactttgagctgactgtcctgccaggcgggctg 1598

Qy 1501 CGCTGNCNWSNGNGARGGNGAYCGCTTYGARYTNACNCTNWSNTGYCARGCNGGNTN 1560

Db 1599 cccaaggaagcctgcattggagatctcatcgccaggtggccagcccttgcacagcggctg 1658

Qy 1561 CCNAAARGCNGTGYATGGARATHWSNWSNCCNGGNTGYCARCCNCCNCGNCARMGNTN 1620









Sequence 1047 BP;

DR WPI; 91-051337/01.

[illegible]

RESULT	7
ID	T05466 standard; cDNA to mRNA; 36 BP.
AC	T05466;
DT	23-JAN-1996 (first entry)
DE	Sequence encoding immunogenic peptide of melanoma antigen gp100.
KW	Melanoma; antigen; vaccine; immunogen; primer; probe; detection;
KW	identification; tumour; gp100; ss.

DE Sequence encoding immunogenic peptide of melanoma antigen gp100.  
KW Melanoma; antigen; vaccine; immunogen; primer; probe; detection;  
KW identification; tumour; gp100; ss.  
OS Homo sapiens.

FH	Key	Location/Qualifiers
FT	CDS	1..36

FT	/tag= b		
FT	protein_bind	7..36	
FT	/tag= c		
FT	protein_bind	7..33	
FT	/tag= d		
FT	protein_bind	10..36	
FT	/tag= e		
PB	EP-668350-A1.		
PD	23-AUG-1995.		
PF	14-FEB-1995; 200348.		
PR	16-FEB-1994; EP-200337.		
PR	21-DEC-1994; EP-203709.		
PA	(ALKU ) AKZO NOBEL NV.		
PI	Adema GJ, Figdor CG;		
DR	WPI; 95-284790/38.		
DR	P-PSDB; R78642.		
PT	Melanoma associated antigen gp100 - used in vaccines and for the		
PS	detection of tumours		
PS	Claim 7; Page 27; 40pp; English.		
CC	Immunogenic peptides derived from the melanoma associated antigen		
CC	(See R78639-45) may be used in the production of vaccines.		
CC	Nucleotide sequences encoding the immunogenic peptides may be used		
CC	as primers and probes in the detection of melanoma cells. Tumour		
CC	infiltrating lymphocytes capable of binding to the melanoma		
CC	associated antigen can be cultured ex vivo and returned to melanoma		
CC	particles, and when radiolabelled, they may be used to identify		
CC	tumour deposits.		
SQ	Sequence 36 BP; 9 A; 8 C; 11 G; 8 T;		
<hr/>			
Query Match		2.0%; Score 150; DB 16; Length 36;	

01--NOV--1989.

```
FT      /label= N-glycos_site
```

01--NOV--1989.

FT	/*tag= a	
FT	/note="Sticky end"	complement (199..202)
FT	misc feature	
FT	/*tag= b	
FT	/note="Sticky end"	
FT	misc feature	1..23
FT	/*tag= c	
FT	/label=Oligo PA9	
FT	misc feature	24..46
FT	/*tag= d	
FT	/label=Oligo PA11	
FT	misc feature	47..74
FT	/*tag= e	
FT	/label=Oligo PA13	
FT	misc feature	75..99
FT	/*tag= f	
FT	/label=Oligo PA15	
FT	misc feature	100..128
FT	/*tag= g	
FT	/label=Oligo PA17	
FT	misc feature	129..150
FT	/*tag= h	
FT	/label=Oligo PA19	
FT	misc feature	151..176
FT	/*tag= i	
FT	/label=Oligo PA21	
FT	misc feature	177..202
FT	/*tag= j	
FT	/label=Oligo PA23	
FT	misc feature	complement (5..27)
FT	/*tag= k	
FT	/label=Oligo PA10	
FT	misc feature	complement (28..50)
FT	/*tag= l	
FT	/label=Oligo PA12	
FT	misc feature	complement (51..78)
FT	/*tag= m	
FT	/label=Oligo PA14	
FT	misc feature	complement (79..103)
FT	/*tag= n	
FT	/label=Oligo PA16	
FT	misc feature	complement (104..132)
FT	/*tag= o	
FT	/label=Oligo PA18	
FT	misc feature	complement (133..154)
FT	/*tag= p	
FT	/label=Oligo PA20	
FT	misc feature	complement (155..180)
FT	/*tag= q	
FT	/label=Oligo PA22	
FT	misc feature	complement (181..202)
FT	/*tag= r	
FT	/label=Oligo PA24	
PN	EP-294979-A.	
PD	14-DEC-1988.	
PD	31-MAY-1988.	304906.
PR	10-JUN-1987; GB-013559.	
PR	(PACO-) PA Consult Services.	
PI	Edwards RM, Light JA, Nicholson K;	
PI	WPI; 88-355406/50.	
DR	Recombinant plasmids contg. DNA region	
PT	produce novel silk-like structural pro	
PT	mori fibroid	
PS	Claim 7; Page 5; 17pp; English.	

PD 14-DEC-1988.  
PF 31-MAY-1988; 304906.  
PR 10-JUN-1987; GB-013559.  
PA (PACO-) PA Consult Services.  
PI Edwards RM, Light JA, Nicholson K;  
DR WPI; 88-355406/50.  
PT Recombinant plasmids contg. DNA region and lac Z gene portion -  
PT produce novel silk-like structural proteins similar to Bombyx  
PT mori fibroid  
PS Claim 7; Page 5; 17pp; English.

ID	Q59596 standard; cDNA; 405 BP.
AC	Q59596;

FT	unsure	41..43
----	--------	--------

FT	/tag= f	"encodes leu"	47..49
FT	/note=	unsure	
FT	/tag= g	"encodes leu"	50..52
FT	/note=	unsure	
FT	/tag= h	"encodes leu"	56..58
FT	/note=	unsure	
FT	/tag= i	"encodes leu"	98..100
FT	/note=	unsure	
FT	/tag= j	"encodes leu"	107..109
FT	/note=	unsure	
FT	/tag= k	"encodes ser"	113..115
FT	/note=	unsure	
FT	/tag= l	"encodes leu"	125..127
FT	/note=	unsure	
FT	/tag= m	"encodes leu"	131..133
FT	/note=	unsure	
FT	/tag= n	"encodes leu"	146..148
FT	/note=	unsure	
FT	/tag= o	"encodes arg"	173..175
FT	/note=	unsure	
FT	/tag= p	"encodes arg"	176..178
FT	/note=	unsure	
FT	/tag= q	"encodes arg"	191..193
FT	/note=	unsure	
FT	/tag= r	"encodes leu"	212..214
FT	/note=	unsure	
FT	/tag= s	"encodes leu"	236..238
FT	/note=	unsure	
FT	/tag= t	"encodes ser"	239..241
FT	/note=	unsure	
FT	/tag= u	"encodes leu"	248..250
FT	/note=	unsure	
FT	/tag= v	"encodes leu"	254..256
FT	/note=	unsure	
FT	/tag= w	"encodes leu"	263..265
FT	/note=	unsure	
FT	/tag= x	"encodes ser"	266..268
FT	/note=	unsure	
FT	/tag= y	"encodes leu"	275..277
FT	/note=	unsure	
FT	/tag= z	"encodes arg"	302..304
FT	/note=	unsure	
FT	/tag= aa		

FT	/note= "encodes Ser"
FT	unsure 311..313
FT	*tag= ab
FT	/note= "encodes Ser"
FT	unsure 314..316
FT	*tag= ac
FT	/note= "encodes Leu"
FT	unsure 323..325
FT	*tag= ad
FT	/note= "encodes Leu"
FT	CDS 11..343
FT	*tag= ae
FT	/product= human_preinsulin
PN	BE-885196-A.
PD	31-DEC-1980.
PF	11-SEP-1980; 885196.
PR	12-SEP-1979; US-075192.
PA	(REGC ) UNIV OF CALIFORNIA.
DR	WPI; 81-05762D/05.
DR	P-PSDB; P10053.
PT	DNA transfer vectors contg. codes for human insulin precursors - used to transform microorganisms for insulin prodn.
PT	Claim 4; Page 33; 50pp; French.
CC	This coding sequence is a generic coding sequence which covers DNA sequences able to code for the human preproinsulin amino acid sequence. The features table indicates where Leu, Ser or Arg codons are intended. These 3 amino acids cannot readily be specified by a single generic codon, hence the need for clarification. For each of CC Leu, Arg and Ser the following pairs of codons are what was CC intended, respectively: YTR and CTY (= Leu), MGR and CGY (= Arg) CC and TON and AGY (= Ser). Transfer vectors containing DNA whose CC sequence corresponds to this generic sequence are claimed.
CC	See also N10052.
SQ	Sequence 416 BP; 59 A; 72 C; 85 G; 59 T;
Query Match 1.7%; Score 128; DB 4; Length 416;	
Best Local Similarity 6.2%; Pred. No. 4.65e+01;	
Matches 10; Conservative 50; Mismatches 102; Indels 0; Gaps 0;	
Db	122 gcnvtatvntgtngcgcargmrgcnttlyttayacncnaaracnmqmgngar 181   :  :::   ::::   ::::   ::::   ::::   ::::   ::::   ::::
Cp	1471 GNACDATRTNARNGTNACWSRAANWSNCCTTANCRTATNARNACRCACTNARNGGA 1412
Db	182 gcngargavtyncargtngcargtngarytngtngdngngccngcngcnwsgnytn 241 :   : : : : : : : : : : : : : : : : : : :
Cp	1411 CYTGCKYTTNACRNCKNARNGTCNCGTNCCTCTNARNARNGGCCNARWNSCCNG 1352 :: : : : : : : : : : : : : : : : : : :
Db	242 carccnytgngyngarggnwnsyntncaraarmgngnath 283 :: : : : : : : : : : : : : : : : : : :
Cp	1351 TDATNWSYTGTGNTWSCATDATNWSNWSNGORTONGNGCCYT 1310
RESULT 15	
ID	Q88233 standard; DNA; 420 BP.
AC	Q88233;
DT	25-SEP-1995 (first entry)
KM	Bovine arGF synthetic gene.
DE	Acidic fibroblast growth factor; aFGF; vulnerary; angiogenesis;
KW	mitogen; ss.
OS	Synthetic.
PN	US5401832-A.
PD	28-MAR-1995.
PF	24-DEC-1984; 685923.
PR	24-DEC-1984; US-685923.



Aug 27 18:59

US-08-231-565A-27.rng

23

```

PR 12-SEP-1985; US-774359.
PR 30-MAY-1986; US-868473.
PR 11-JUL-1986; US-884460.
PR 04-JUN-1987; US-054991.
PR 04-MAY-1988; US-190293.
PR 08-FEB-1991; US-654397.
PR 25-SEP-1991; US-765472.
PR 25-SEP-1992; US-951365.
PA (MERI ) MERCK & CO INC.
PI Gimenez-Callego G, Kelly LJ, Linemeyer DL, Thomas KA;
DR WPI; 95-I38983/18.
DR P-PSDB; R74648.
PT New recombinant human acidic fibroblast growth factor - used to
PT promote cell growth, to promote wound healing, for vascular
PT grafts and blood vessel repair
PS Disclosure; Table 1; 25pp; English.
CC Oligonucleotides were synthesized on the basis of the amino
CC acid sequence of cattle acidic fibroblast growth factor (aFGF)
CC and used to construct a synthetic gene (given in Q88233) that
CC incorporates codons preferred by E. coli or mammalian cells,
CC unique cloning sites, etc. The preferred sequence is given in
CC Q88234. The synthetic gene can be mutated to obtain a human aFGF
CC synthetic gene for production of recombinant protein (R74647).
SQ Sequence 420 BP; 87 A; 37 C; 62 G; 55 T;

Query Match 1.7%; Score 128; DB 14; Length 420;
Best Local Similarity 6.8%; Pred. No. 4.65e+01;
Matches 12; Conservative 63; Mismatches 102; Indels 0; Gaps 0;

Db 178 garacngncarttytngcncatggaacngagngnyntntayggngwancarcncn 237
: : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Cp 1966 RNANGGWSRTTYTNCDDATNGRCANWSRCARADATNCKNGNARCNCKNCKRT 1907
: : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 238 aaygargartgytnttytngarmgnytnrgaraaaycaytayaacntayathwn 297
: : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Cp 1906 GNWSNWSNWSRTGNGGNARYTNGGNACNWSRAARTCYTTCATNARNCKNCKNCKRT 1847
: : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 298 aaraaracaygcngaraacaytggttygtngnytnaaraaayggmgngwnaear 354
: : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Cp 1846 ADATNARNWSNCCNARNACNACGCCATNARNACNARNARDATNCCNACDATNARNG 1790

```

Search completed: Tue Aug 27 19:11:30 1996  
Job time : 172 secs.

\*\*\*\*\*

WATERMAN

(TM)

\*\*\*\*\*

Release 2.1D John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_tpn n.a. - n.a. Smith-Waterman search, using a protein query  
which has been backtranslated into n.a. using IUPAC symbols

Run on: Tue Aug 27 19:11:52 1996; MaePar time 767.73 Seconds  
917.837 Million cell updates/sec

Tabular output not generated.

Title: &gt;US-08-231-565A-27

Description: (1-661) from US08231565A.pep

Perfect Score: 7622

N.A. Sequence:

Comp: 1 ATGCAYYNGTNTNARMG.....TNTYNSNGNCARCAGTN 1983

TACCTRRANCANRANTTYK.....ANRANSWNCNGTGYTCAN

Scoring table:

Gap 30

Nmatch STD : Dbase 0; Query 0

Searched: 509049 seqs, 177673129 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

EST-STS  
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8  
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14  
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20  
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26  
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32  
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38  
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44  
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50  
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56  
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62  
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68  
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74  
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80  
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86  
87:EST87 88:EST88 89:EST89 90:EST90 91:STS1 92:STS2  
93:STS3 94:STS4 95:STS5 96:STS6 97:STS7  
EST-STS-TWO  
98:EST91 99:EST92 100:EST93 101:EST94 102:EST95  
103:EST96 104:EST97 105:EST98 106:EST99 107:EST100  
108:EST101 109:gnEST1 110:gnEST2 111:gnEST3 112:gnEST4  
113:gnEST5 114:gnEST6 115:gnEST7 116:gnEST8 117:gnEST9  
118:gnEST10 119:gnEST11 120:gnEST12 121:gnEST13  
122:gnEST14 123:gnEST15 124:gnSTS1 125:gnSTS2 126:gnEST1

Database:

93:STS3 94:STS4 95:STS5 96:STS6 97:STS7

EST-STS-TWO

98:EST91 99:EST92 100:EST93 101:EST94 102:EST95

103:EST96 104:EST97 105:EST98 106:EST99 107:EST100

108:EST101 109:gnEST1 110:gnEST2 111:gnEST3 112:gnEST4

113:gnEST5 114:gnEST6 115:gnEST7 116:gnEST8 117:gnEST9

118:gnEST10 119:gnEST11 120:gnEST12 121:gnEST13

122:gnEST14 123:gnEST15 124:gnSTS1 125:gnSTS2 126:gnEST1

127:enEST2 128:enEST3 129:enEST4 130:enEST5 131:enEST6  
132:enEST7 133:enEST8 134:enEST9 135:enEST10 136:enEST11

Statistics: Mean 69.416; Variance 71.291; scale 0.974

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	1939	25.4	621	114	N28728	y67a12.r1 Homo sapie	0.00e+00
2	1939	25.4	621	136	HS728262	y67a12.r1 Homo sapie	0.00e+00
3	1812	23.8	638	114	N28736	y67c12.r1 Homo sapie	0.00e+00
4	1812	23.8	638	136	HS736264	y67c12.r1 Homo sapie	0.00e+00
5	1510	19.8	546	136	HS746262	y67f11.r1 Homo sapie	0.00e+00
6	1510	19.8	546	114	N28746	y67f11.r1 Homo sapie	0.00e+00
7	1350	17.7	521	118	N40290	y80c07.r1 Homo sapie	0.00e+00
8	1350	17.7	521	131	HS290283	y80c07.r1 Homo sapie	0.00e+00
9	1297	17.0	416	68	R78477	y13d06.r1 Homo sapie	0.00e+00
10	1006	13.2	306	83	T29726	EST92123 Homo sapiens	0.00e+00
11	974	12.8	286	46	R01509	ye79a02.r1 Homo sapie	3.85e-302
12	948	12.4	430	113	N23118	y67c12.s1 Homo sapie	6.28e-292
13	948	12.4	430	130	HS118284	y67c12.s1 Homo sapie	6.28e-292
14	877	11.5	431	113	N23111	y67a12.s1 Homo sapie	3.75e-264
15	877	11.5	431	129	HS111256	y67a12.s1 Homo sapie	3.75e-264
16	292	3.8	506	130	HS123264	y47e01.r1 Homo sapie	1.70e-48
17	292	3.8	506	114	N29123	y47e01.r1 Homo sapie	1.70e-48
18	173	2.3	495	19	HS1731	yp85g02.r1 Homo sapie	3.27e-13
19	171	2.2	58	130	HS128259	y67f11.s1 Homo sapie	1.07e-12
20	171	2.2	58	113	N23128	y67f11.s1 Homo sapie	1.07e-12
21	154	2.0	522	81	T22039	4047 Arabidopsis thal	1.66e-08
22	147	1.9	336	84	T34583	EST1046 Homo sapiens	7.10e-07
23	144	1.9	493	113	N24647	y89g12.s1 Homo sapie	3.40e-06
24	144	1.9	493	135	HS647262	y89g12.s1 Homo sapie	3.40e-06
25	138	1.8	127	5	H03700	yj42f09.r1 Homo sapie	7.15e-05
26	140	1.8	288	83	T31913	EST40905 Homo sapiens	2.62e-05
27	138	1.8	306	93	G05045	human STS WI-6432	7.15e-05
28	137	1.8	338	40	HSDEK123	H. sapiens partial cD	1.17e-04
29	134	1.8	354	35	HSC17H061	H. sapiens partial cD	5.10e-04
30	141	1.8	360	121	CEIK087C2F	C.elegans cDNA clone	1.58e-05
31	141	1.8	360	127	CEK087C2F	C.elegans cDNA clone	1.58e-05
32	134	1.8	364	108	T99127	ye62d08.r1 Homo sapie	5.10e-04
33	139	1.8	389	6	H10164	ym05e09.s1 Homo sapie	4.34e-05
34	139	1.8	442	55	R36012	y69a03.r1 Homo sapie	4.34e-05
35	134	1.8	450	87	T44050	7313 Arabidopsis thal	5.10e-05
36	138	1.8	495	116	N34300	yy51f12.s1 Homo sapie	7.15e-05
37	138	1.8	495	131	HS300270	yy51f12.s1 Homo sapie	7.15e-05
38	136	1.8	498	47	R04666	pk27b11.s1 Caenorhabd	1.92e-04
39	138	1.8	519	9	H18327	yn49a07.s1 Homo sapie	7.15e-05
40	133	1.7	192	12	H26106	y144f12.s1 Homo sapie	8.25e-04
41	133	1.7	371	108	T97699	ye54c03.s1 Homo sapie	8.25e-04
42	133	1.7	383	93	G07853	human STS CHIC.CAVA72	8.25e-04
43	132	1.7	431	27	H78737	yu49f07.r1 Homo sapie	1.33e-03
44	132	1.7	511	98	T62611	yc04c12.r1 Homo sapie	1.33e-03
45	133	1.7	543	8	HS5883	ym22e11.r1 Homo sapie	8.25e-04

## ALIGNMENTS

RESULT 1







vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B

Qy 142

[illegible][illegible]

6	RESULT	
	LOCUS	N28746 546 bp mRNA EST 04-JAN-1996
	DEFINITION	Y87f11.r1 Homo sapiens cDNA clone Z66829 5' similar to SP:PIR:A53668 A53668 glycoprotein gp100 precursor, melanocyte lineage - human ;.
	ACCESSION	N28746
	NTID	g1146982
	KEYWORDS	EST.
	SOURCE	human clone=Z66829 primer=T7 library=Seares melanocyte 2NBHM vector=T7T3D (Pharmacia) with a modified polylinker host=DH10B







```

NCBI gi: 611824
Location/Qualifiers
source
  1..306
    /organism="Homo sapiens"
    /note="human"
mRNA
BASE COUNT      63 a   77 c   86 g   75 t   5 others
ORIGIN

Query Match      13.2%; Score 1006; DB 83; Length 306;
Best Local Similarity 54.4%; Pred. No. 0.00e+00;

```

TITLE	COMMENT
The WashU-Merck EST Project Unpublished (1995)	

Contact: Wilson EK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
High quality sequence starts: 1  
High quality sequence stops: 1  
Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL ; contact  
IMAGE Consortium (info@image.llnl.gov) for further information

Trace considered overall poor quality  
Possible reversed clone: similarity on wrong strand.

NCBI gi: 751245  
Location/Qualifiers  
1..286  
/organism="Homo sapiens"  
/clone="123914"  
/note="human"

BASE COUNT 50 a 86 c 87 g 59 t 4 others  
ORIGIN

Query Match 12.8%; Score 974; DB 46; Length 286;  
Best Local Similarity 51.8%; Pred. No. 3.85e-302;  
Matches 147; Conservative 67; Mismatches 70; Indels 0; Gaps 0;

Db 1 ttccgtccaccctgacattgtcnggaattgaaagtgcagagatcctgcaggtgtgcc 60  
::: || || : || || || : || || || : || || || : || || || : || || ||  
Qy 1449 YWNGTACNTGATATGTCNARGGNATHGAPWSGNGCARATHYTCARGCGTNC 1508  
::: || || : || || || : || || || : || || || : || || || : || || ||  
Db 61 tncctgaggggagtgatgctgctgctgccaangcggtgcaccaagga 120  
: : || || || || || || || || || || : || || || : || || || : || || ||  
Qy 1509 NWSNGNGARGGNGGAYGCTTGARYTNACNGTNWSNTGYCARGNGGNTNCCNAARGA 1568  
: : || || || || || || || || || || : || || || : || || || : || || ||  
Db 121 agcctgcatgagatctatcgcagaggtgcccagccctgcgcagcgtgtgcagcct 180  
: || || || || || || || || || || : || || || : || || || : || || ||  
Qy 1569 RGCTGTGATGARATHWSNCCNGGNTGYCARCNCNGCNCARMGNTYNTGYCARCC 1628  
: || || || || || || || || || || : || || || : || || || : || || ||  
Db 181 ggtgtcancaccagcagctgctgctgctgacacagatactgaaggtggtcgtcggg 240  
: || : || : || || || || || || : || || || : || || || : || || ||  
Qy 1629 NGTYNINCCNSNCNGTGYCARYTNGTNTNCAVCARATHYTNARGGNGWSNGG 1688  
: || : || : || || || || || || : || || || : || || || : || || ||  
Db 241 gacatactcctcaatgtctctgctgataccacagcctgg 284  
|| || || || || || || : || || || || || || || || || : || ||  
Qy 1689 NAGNTATGCTYTNAAAGTWSNTNGCNGAYACNAAYWSNYTNG 1732  
|| || || || || || || : || || || || || || || || || : || ||

RESULT 12  
LOCUS N23118 430 bp mRNA EST 28-DEC-1995  
DEFINITION yx67c12.s1 Homo sapiens cDNA clone 266806 3' similar to SP:A41234  
A41234 MELANOCYTE-SPECIFIC PROTEIN PMEL-17 PRECURSOR - ;  
ACCESSION N23118  
NID g1137268  
KEYWORDS EST.  
SOURCE human clones=266806 primer=m13 -40 forward library=Soares melanocyte  
2Nblm vector=p7T3D (Pharmacia) with a modified polylinker  
host=DH10B (ampicillin resistant) RseI=Not I RseI=2-Eco RI Male.  
1st strand cDNA was primed with a Not I - oligo(dT) primer  
[5'-TGTTACCAATCTGAAGTGGAGCGCGGAGTGTGTTTTTTTTTTTTTTT-3'],  
double-stranded cDNA was size selected, ligated to Eco RI adapters  
(Pharmacia), digested with Not I and cloned into the Not I and Eco  
RI sites of a modified pT73 vector (Pharmacia). Library  
constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal  
fore skin melanocytes (FS374) was kindly provided by Dr. Anthony P.  
Albino.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;  
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;  
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 430)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
Parsons, J., Rifkin, L., Rohlffing, T., Soares, M., Tan, F.,

## REFERENCE

DE yx67c12.s1 Homo sapiens cDNA clone 266806 3' similar to SP:A41234  
DE A41234 MELANOCYTE-SPECIFIC PROTEIN PMEL-17 PRECURSOR - ;  
KW EST.

Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and  
Wilson, R.  
The WashU-Merck EST Project  
Unpublished (1995)

## TITLE

## JOURNAL

## COMMENT

Contact: Willson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu  
High quality sequence stops: 288  
Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.

NCBI gi: 1137268  
Location/Qualifiers  
1..430  
/organism="Homo sapiens"  
/clone="266806"  
/note="human"  
<1..>430

BASE COUNT 114 a 107 c 115 g 93 t 1 others  
ORIGIN

Query Match 12.4%; Score 948; DB 113; Length 430;  
Best Local Similarity 53.1%; Pred. No. 6.28e-292;  
Matches 169; Conservative 56; Mismatches 92; Indels 1; Gaps 1;

Db 113 acctgctgccactgagggggggtctattctcacaatgggacagagcagaagatg 172  
|| || || || : || : || : || || || || || || : || || || : || || ||  
Cp 1982 ACYTGTCGCCNWSNARNRNGWSRTTYTCNCDDATNGRCANWSRCARAADATNCKN 1923  
|| || || || : || : || : || || || || || || : || || || : || || ||  
Db 173 ggtagacgcagcagtgactgctgctgctggaactgggtgaggaagtcttcttc 232  
|| || : || || || || : || || || || || || || : || || || : || || ||  
Cp 1922 GGNARNCKNARCARTGNWSNWSNWSRTGNGNARYTNGNACNWSRARTCYTCTTC 1863  
|| || : || || || || : || || || || || || || : || || || : || || ||  
Db 233 ataagtctgcgcctatatcatagatgcaagaccacagccatcaaccagcaagatg 292  
|| || : || : || || || || || : || || || || || || : || || || : || || ||  
Cp 1862 ATNARNCKNCKCKRTADATNARNWSNCGNARNACNACGCCATNARNARNADATN 1803  
|| || : || : || || || || || : || || || || || || : || || || : || || ||  
Db 293 cccacgacagcaggaactgcccagggcgtctcttgaccagggcatgaagctgggtg 352  
|| || || || : || || || || || : || || || || || || : || || || : || || ||  
Cp 1802 CCNACDATTNARNGNACYTGCNARNCCGCTCYTGCNCCGCCATDATNARYTGNTN 1743  
|| || || || : || || || || || : || || || || || || : || || || : || || ||  
Db 353 cttgacacactgccaggtcttgatcagcagacacacattgagcagtgatgtcccca 412  
: || || || : || || || || || : || || || || || || : || || || : || || ||  
Cp 1742 -WSNACNACNARNWSRTYNGTTCNCGNARNWSNACRTNARNRCARTANGTNCNWS 1684  
|| || || || : || || || || || : || || || || || || : || || || : || || ||  
Db 413 gccacccttcagatctg 430  
|| || || || : || || || || || : || || || || || || : || || || : || || ||  
Cp 1683 NCCNCTTTNARDATYTG 1666  
|| || || || : || || || || || : || || || || || || : || || || : || || ||

RESULT 13  
ID HS118284 standard; RNA; EST; 430 BP.  
AC N23118;

DT 29-DEC-1995 (Rel. 46, Created)  
DT 29-DEC-1995 (Rel. 46, Last updated, Version 1)

DE yx67c12.s1 Homo sapiens cDNA clone 266806 3' similar to SP:A41234  
DE A41234 MELANOCYTE-SPECIFIC PROTEIN PMEL-17 PRECURSOR - ;  
KW EST.



Cp 1802 CCNACDATTNRAGNACYTCGCCNARNCCN-CCYCTCYTCGNCNGCCATDATTNRYGCT 1744  
 Db 353 gctgaccacttgcaggctgttggtatcagccagagacacatttgaggcagtatgtgcc 412  
 Cp 1743 NWSNACMACN-CCNARNWSRTGTGTCGNCNARNWSNACRTTN-ARCAATTANGTNCN 1686  
 Db 413 gagccacccttcagtatt 430  
 Cp 1685 WSNCCNCCYTTNARDY 1668

```

RESULT 15
AC ID HS111256 standard; RNA; EST; 431 BP.
AD N23111;
DT 29-DEC-1995 (Rel. 46, Created)
DE yx67ai2.s1 Homo sapiens cDNA clone 266782.3' similar to SP:A41234
DE A41234 MELANOCYTE-SPECIFIC PROTEIN PMEL-17 PRECURSOR - ;
KW EST.
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
NC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
RN [1]
RP 1-431
RR Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA Trevasis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;
RT "The WashU-Merck EST Project";
RL Unpublished.
CC Contact: Wilson RK WashU-Merck EST Project Washington University
CC School of Medicine 4404 Forest Park Parkway, Box 8501, St. Louis,
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC est@watson.wustl.edu High quality sequence stops: 286 Source: IMAGE
CC Consortium, LLNL This clone is available royalty-free through LLNL
CC i contact the IMAGE Consortium (info@image.llnl.gov) for further
CC information. NCBI gi: 1137261
FF Key Location/Qualifiers
FF source 1..431
FT /organism="Homo sapiens"
FT /clone="266782"
FT /note="human"
FF mRNA <1..>431
FF Sequence 431 BP; 114 A; 105 C; 116 G; 95 T; 1 other;

Query Match 11.5%; Score 877; DB 129; Length 431;
Best Local Similarity 52.8%; Pred. No. 3.75e-264;
Matches 168; Conservative 56; Mismatches 91; Indels 3; Gaps
113 acctgctgccacgcagtgaggggcgattctcaccaatgggaagaagcagaagtcggc 172
|||:||| || | : : : |||:|||:|||:||| :|||:||| :
Cp 1982 ACYGTGTCNCSWNNARNNGNSRITTYTCCDATTNGRCANWSRCARADAATCNKN 1923
|||:||| || | : : : |||:|||:|||:||| :|||:||| :
Db 173 gttagacgcagccagtactgctatgttgcacatgggtagcaggaaagtcttgttc 232
|||:||| ||| : : : |||:||| :|||:||| ||| :|||:|||:|||
Cp 1922 GGNARNCKNARCARTGNWSNWSNRSTNGNARYTGNGNAGNCSRAARTCTGTTC 1863
|||:||| ||| : : : |||:||| :|||:||| ||| :|||:|||:|||
Db 233 ataagtctgcgcctatatcatagagatgcaaggaccacagccatcaaccaccagaagt 292
|||:||| ||| : : : |||:||| :|||:||| ||| :|||:|||:|||
Cp 1862 ATARNCKNCKKCRATDATNARNNSGNCNARNACNCGCCATNARNACNARNARDATN 1803
|||:||| ||| : : : |||:||| :|||:||| ||| :|||:|||:|||
Db 293 ccacagatcagcggaaacctgcccaaggccttgccttgcaccagcgatgataagctgggt 352
|||:||| ||| : : : |||:||| :|||:||| ||| :|||:|||:|||

```

Query Match	11.5%;	Score 877;	DB 129;	Length 431;
Best Local Similarity	52.8%;	Pred. No. 3,75e-264;		
Matches	168;	Conservative 56;	Mismatches 91;	Indels 3; Gaps 3;
Ddb	113	acctgctgccactgaggagggcggtattccaccaataggacaagcagaagaatgcgg	172	:
Cp	1982	ACYCTGTGNCNWNARNNGNWRSRTTYTCCDATTGGRCANWSRCAADATNCKN	1923	:
Ddb	173	qgtagacgaqcacagtgactgctgctatgtggcaactggggtacggagaagcttctgcttc	232	:
Cp	1922	CGNARCNKARCARTGNWNSWNRSGTGNNARYTGGNGACNWSRAARTCYGTTC	1863	:
Ddb	233	ataagctctgcgcctatatatcacagatgcaagaccacagcccatcaacaccagaagaatg	292	:
Cp	1862	ATNARNCKNKCKTKRTADATNARNWSNGCNAARNACNAGCCCATNARNACVARNARDATN	1803	:
Ddb	293	cccacgatcagcggaacctgcccaaggccttgcttttgaccaggcatgatgaagctgggt	352	:

=> d his

(FILE 'USPAT' ENTERED AT 10:53:45 ON 27 AUG 96)

L1 155 S MART OR GP100  
L2 7055 S NUCLEIC ACID  
L3 4343 S CDNA  
L4 8911 S L2 OR L3  
L5 14 S L1 AND L4  
L6 7 S L1(L)L4  
L7 0 S L1(P)L4  
L8 7 S L6

=> d 1-7 cit

1. ~~5,484,891~~, Jan. 16, 1996, Selectin ligands; Laurence A. Lasky, et al., 530/387.3; 435/7.2; 530/350, 395 [IMAGE AVAILABLE]

2. ~~5,464,815~~, Nov. 7, 1995, Inhibition of heparin-binding; Steven Chamow, et al., 514/8; 424/85.2; 436/86, 87; 514/21; 530/412 [IMAGE AVAILABLE]

3. ~~5,369,025~~, Nov. 29, 1994, Recombinant fowlpox vaccine for protection against Marek's disease; Keyvan Nazerian, et al., 435/235.1; 424/186.1, 199.1, 229.1; 435/69.3, 172.3, 240.2; 530/350; 536/23.72; 935/9, 32, 34, 57, 63, 65, 70 [IMAGE AVAILABLE]

4. ~~5,318,890~~, Jun. 7, 1994, Assays for inhibitors of leukocyte adhesion; Steven Rosen, et al., 435/7.24, 7.1, 7.2, 7.92; 530/387.3 [IMAGE AVAILABLE]

5. ~~5,304,640~~, Apr. 19, 1994, DNA sequence encoding a selectin ligand; Laurence A. Lasky, et al., 536/23.5; 435/69.1, 172.3, 240.2, 320.1 [IMAGE AVAILABLE]

6. ~~5,242,829~~, Sep. 7, 1993, Recombinant pseudorabies virus; Dennis L. Panicali, et al., 435/320.1; 424/199.1, 229.1, 232.1; 435/69.1, 69.3, 172.3 [IMAGE AVAILABLE]

7. ~~5,141,922~~, Aug. 25, 1992, Biologically active proteins and a method for their use; Gwen G. Krivi, 514/12, 21; 530/324, 399 [IMAGE AVAILABLE]

=>

=> d his

(FILE 'USPAT' ENTERED AT 10:46:09 ON 27 AUG 96)

E KAWAKAMI, Y/IN

E ROSENBERG, S/IN

L1 24 S E15,E14  
L2 0 S MELANONA(3A)ANTIGEN#(5A)(LYMPHOCYTE# OR LYMPHOKYTE#)  
L3 3 S MELANOMA(3A)ANTIGEN#(5A)(LYMPHOCYTE# OR LYMPHOKYTE#)  
L4 7055 S NUCLEIC ACID  
L5 4343 S CDNA  
L6 8911 S L4 OR L5  
L7 0 S L6 AND L3  
L8 9 S L1 AND L6  
L9 0 S L3 AND L8

=>